Université de Montréal

Bootstrapping high frequency data

par
Koomla Ulrich Hounyo

Département de sciences économiques
Faculté des arts et des sciences

Thèse présentée à la Faculté des arts et des sciences
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences économiques

Août, 2013

Université de Montréal
Faculté des arts et des sciences

Cette thèse intitulée :

Bootstrapping high frequency data

présentée par :

Koomla Ulrich Hounyo

a été évaluée par un jury composé des personnes suivantes :

Benoit Perron,	président-rapporteur
Silvia Gonçalves,	directrice de recherche
Ilze Kalnina,	codirectrice de recherche
Marine Carrasco,	membre du jury
Russell Davidson,	examinateur externe
Pierre Duchesne,	représentant du doyen de la FAS

Thèse acceptée le

Résumé

Nous développons dans cette thèse, des méthodes de bootstrap pour les données financières de hautes fréquences. Les deux premiers essais focalisent sur les méthodes de bootstrap appliquées à l'approche de "pré-moyennement" et robustes à la présence d'erreurs de microstructure. Le "pré-moyennement" permet de réduire l'influence de l'effet de microstructure avant d'appliquer la volatilité réalisée. En se basant sur cette approche d'estimation de la volatilité intégrée en présence d'erreurs de microstructure, nous développons plusieurs méthodes de bootstrap qui préservent la structure de dépendance et l'hétérogénéité dans la moyenne des données originelles. Le troisième essai développe une méthode de bootstrap sous l'hypothèse de Gaussianité locale des données financières de hautes fréquences.

Le premier chapitre est intitulé: *"Bootstrap inference for pre-averaged realized volatility based on non-overlapping returns"*. Nous proposons dans ce chapitre, des méthodes de bootstrap robustes à la présence d'erreurs de microstructure. Particulièrement nous nous sommes focalisés sur la volatilité réalisée utilisant des rendements "pré-moyennés" proposés par Podolskij et Vetter (2009), où les rendements "pré-moyennés" sont construits sur des blocs de rendements à hautes fréquences consécutifs qui ne se chevauchent pas. Le "pré-moyennement" permet de réduire l'influence de l'effet de microstructure avant d'appliquer la volatilité réalisée. Le non-chevauchement des blocs fait que les rendements "pré-moyennés" sont asymptotiquement indépendants, mais possiblement hétéroscédastiques. Ce qui motive l'application du wild bootstrap dans ce contexte. Nous montrons la validité théorique du bootstrap pour construire des intervalles de type percentile et percentile-t. Les simulations Monte Carlo montrent que le bootstrap peut améliorer les propriétés en échantillon fini de l'estimateur de la volatilité intégrée par rapport aux résultats asymptotiques, pourvu que le choix de la variable externe soit fait de façon appropriée. Nous illustrons ces méthodes en utilisant des données financières réelles.

Le deuxième chapitre est intitulé : *"Bootstrapping pre-averaged realized volatility under market microstructure noise"*. Nous développons dans ce chapitre une méthode de bootstrap par bloc basée sur l'approche "pré-moyennement" de Jacod et al. (2009), où les rendements "pré-moyennés" sont construits sur des blocs de rendements à haute fréquences consécutifs qui se chevauchent. Le chevauchement des blocs induit une forte dépendance dans la structure des rendements "pré-moyennés". En effet les rendements "pré-moyennés" sont m-dépendant avec m qui croît à une vitesse plus faible que la taille d'échantillon n. Ceci motive l'application d'un bootstrap par bloc spécifique. Nous montrons que le bloc bootstrap suggéré par Bühlmann et Künsch (1995) n'est valide que lorsque la volatilité est constante. Ceci est dû à l'hétérogénéité dans la moyenne des rendements "pré-moyennés" au carré lorsque la volatilité est stochastique. Nous proposons donc une nouvelle procédure de bootstrap qui combine le wild bootstrap et le bootstrap par bloc, de telle sorte que la dépendance sérielle des rendements "pré-moyennés" est préservée à l'intérieur des blocs et la condition d'homogénéité nécessaire pour la validité du bootstrap est respectée. Sous des conditions de taille de bloc, nous montrons que cette méthode est convergente. Les simulations Monte Carlo montrent que le bootstrap améliore les propriétés en échantillon fini de l'estimateur de la volatilité intégrée par rapport aux résultats asymptotiques. Nous illustrons cette méthode en utilisant des données financières réelles.

Le troisième chapitre est intitulé: *"Bootstrapping realized covolatility measures under local Gaussianity assumption"*. Dans ce chapitre nous montrons, comment et dans quelle mesure on peut approximer les distributions des estimateurs de mesures de co-volatilité sous l'hypothèse de Gaussianité locale des rendements. En particulier nous proposons une nouvelle méthode de bootstrap sous ces hypothèses. Nous nous sommes focalisés sur la volatilité réalisée et sur le beta réalisé. Nous montrons que la nouvelle méthode de bootstrap appliquée au beta réalisé était capable de répliquer les cummulants au deuxième ordre, tandis qu'il procurait une amélioration au troisième degré lorsqu'elle est appliquée à la volatilité réalisée. Ces résultats améliorent donc les résultats existants dans cette littérature, notamment ceux de Gonçalves et Meddahi (2009) et de Dovonon, Gonçalves et Meddahi (2013). Les simulations Monte Carlo montrent que le bootstrap améliore les propriétés en échantillon fini de l'estimateur de la volatilité intégrée par rapport aux résultats asymptotiques et les résultats de bootstrap existants. Nous illustrons cette méthode en utilisant des données financières réelles.

Mots clés: Données haute fréquence, volatilité réalisée, bruit de microstructure, pré-moyennement, betas réalisé, bootstrap, expansions d'Edgeworth.

Abstract

We develop in this thesis bootstrap methods for high frequency financial data. The first two chapters focalise on bootstrap methods for the "pre-averaging" approach, which is robust to the presence of market microstructure effects. The main idea underlying this approach is that we can reduce the impact of the noise by pre-averaging high frequency returns that are possibly contaminated with market microstructure noise before applying a realized volatility-like statistic. Based on this approach, we develop several bootstrap methods, which preserve the dependence structure and the heterogeneity in the mean of the original data. The third chapter shows how and to what extent the local Gaussianity assumption can be explored to generate a bootstrap approximation for covolatility measures.

The first chapter is entitled *"Bootstrap inference for pre-averaged realized volatility based on non-overlapping returns"*. The main contribution of this chapter is to propose bootstrap methods for realized volatility-like estimators defined on pre-averaged returns. In particular, we focus on the pre-averaged realized volatility estimator proposed by Podolskij and Vetter (2009). This statistic can be written (up to a bias correction term) as the (scaled) sum of squared pre-averaged returns, where the pre-averaging is done over all possible non-overlapping blocks of consecutive observations. Pre-averaging reduces the influence of the noise and allows for realized volatility estimation on the pre-averaged returns. The non-overlapping nature of the pre-averaged returns implies that these are asymptotically independent, but possibly heteroskedastic. This motivates the application of the wild bootstrap in this context. We provide a proof of the first order asymptotic validity of this method for percentile and percentile-t intervals. Our Monte Carlo simulations show that the wild bootstrap can improve the finite sample properties of the existing first order asymptotic theory provided we choose the external random variable appropriately.

The second chapter is entitled *"Bootstrapping pre-averaged realized volatility under market microstructure noise "*. In this chapter we propose a bootstrap method for inference on integrated volatility based on the pre-averaging approach of Jacod et al. (2009), where the pre-averaging is done over all possible overlapping blocks of consecutive ob-

servations. The overlapping nature of the pre-averaged returns implies that these are m-dependent with m growing slowly with the sample size n. This motivates the application of a blockwise bootstrap method. We show that the “blocks of blocks” bootstrap method suggested by Politis and Romano (1992) (and further studied by Bühlmann and Künsch (1995)) is valid only when volatility is constant. The failure of the blocks of blocks bootstrap is due to the heterogeneity of the squared pre-averaged returns when volatility is stochastic. To preserve both the dependence and the heterogeneity of squared pre-averaged returns, we propose a novel procedure that combines the wild bootstrap with the blocks of blocks bootstrap. We provide a proof of the first order asymptotic validity of this method for percentile intervals. Our Monte Carlo simulations show that the wild blocks of blocks bootstrap improves the finite sample properties of the existing first order asymptotic theory.

The third chapter is entitled *"Bootstrapping realized volatility and realized beta under a local Gaussianity assumption"*. The financial econometric of high frequency data literature often assumed a local constancy of volatility and the Gaussianity properties of high frequency returns in order to carry out inference. In this chapter, we show how and to what extent the local Gaussianity assumption can be explored to generate a bootstrap approximation. We show the first-order asymptotic validity of the new wild bootstrap method, which uses the conditional local normality properties of financial high frequency returns. In addition to that we use Edgeworth expansions and Monte Carlo simulations to compare the accuracy of the bootstrap with other existing approaches. It is shown that at second order, the new wild bootstrap matches the cumulants of realized betas-based t-statistics, whereas it provides a third-order asymptotic refinement for realized volatility. Monte Carlo simulations suggest that our new wild bootstrap methods improve upon the first-order asymptotic theory in finite samples and outperform the existing bootstrap methods for realized covolatility measures. We use empirical work to illustrate its uses in practice.

Keywords: High frequency data, realized volatility, market microstructure noise, pre-averaging, realized betas, bootstrap, Edgeworth expansions.

Contents

List of Tables

List of Figures

To my family.

Acknowledgments

This thesis would not have been possible without the support of many people. I am espicially grateful to my advisor Professor Silvia Gonçalves for her continuing guidance and encouragement. She taught me how to do research in Econometrics. I am thankful to my co-adviser Ilze Kalnina for valuable discussions and comments.

I would like to thank my co-author Professor Nour Meddahi for useful discussions, constant support and encouragement. I am thankful to my committee members, who offered guidance and support.

My life as a graduate student could not have the same without my friends at Université de Montréal, I would like to thank Arsene Sabas and my classmates Neree Noumon, and Maxime Agbo for their support and continuing friendship throughout these five years.

I thank my parents for their love and support. They always believed in me. No words will ever be enough to describe my gratitude towards them. I thank my wife Berenice for her love and support.

I gratefully acknowledge financial support from CIREQ and Economics department of Université de Montréal.

Introduction Générale

Cette thèse est constituée de trois essais portant sur des méthodes de bootstrap conçues pour les données financières de hautes fréquences. Tout d'abord dans les deux premiers chapitres nous développons des méthodes de bootstrap robustes à la présence d'erreurs de microstructure. Tandis que dans le troisième chapitre nous intéressons à l'hypothèse de Gaussianité locale des données financières de hautes fréquences et nous montrons comment et dans quelle mesure le bootstrap est valide. La volatilité des rendements des actifs financiers est un input indispensable dans les modèles de gestion de risque, sélection de portefeuille, d'actifs dérivés etc. Malheureusement, la volatilite n'est pas directement observable, son estimation avec une précision accrue est donc statistiquement et financièrement d'une grande importance. Plusieurs chercheurs se sont intéressés à comment estimer la volatilité intégrée. En effet la disponibilité de plus en plus grandissante des données à haute fréquence a révolutionné ces dernières décennies le domaine de l'économétrie financière utilisant ces données. Très tôt la volatilité réalisée qui est la somme des carrés des rendements intra-période est devenue très populaire comme étant un estimateur non paramétrique de la volatilite intégrée (la variation quadratique du processus de prix pendant une période fixe, une journée par exemple). L'observation des prix des actifs sans erreurs est l'une des hypothèses fondamentales sous laquelle la volatilité réalisée est un estimateur convergent de la volatilité intégrée. Malheureusement un problème inhérent aux prix à haute fréquence est qu'ils sont contaminés par des frictions dues au marché. En effet, les actifs sur le marché ne sont pas transigés à leur valeur fondamentale. Les frictions incluent les couts de transaction ainsi que l'asymétrie d'information sur le marché. Plusieurs autres facteurs contribuent également à un écart significatif entre le prix efficient et le prix observé par les économètres, parmi lesquels la discrétisation des prix, les erreurs d'arrondi. Tout ce qui contribue à l'écart entre le prix efficient et le prix observé est appelé bruit ou effet de microstructure. En présence d'effet de microstructure, la volatilité réalisée devient un estimateur non convergent de la volatilité intégrée. Plusieurs chercheurs ont proposé de nouveaux estimateurs alternatifs à la volatilité réalisée mais robustes aux effets de microstructure. En présence de bruit de microstructure, nous pouvons utiliser par exemple l'approche par subsampling proposée par Zhang et al. (2005). Elle consiste à combiner deux échelles de temps: haute fréquence et basse fréquence pour obtenir un estimateur convergent de la volatilité intégrée. Une seconde approche développée par Barndorff-Nielsen et al. (2008) utilise une combinaison linéaire des auto-covariances des rendements: c'est le kernel réalisé, il converge avec la vitesse optimale. Enfin, la troisième approche est celle du pré- moyennement, introduite par Podolskij et Vetter (2009) et ensuite généralisée par Jacod et al. (2009). Le "pré-moyennement" permet de réduire l'influence de l'effet de microstructure avant d'appliquer la volatilité réalisée. Cette approche permet un meilleur traitement des effets de bord et est valide sous des hypothèses moins contraignantes sur le bruit. Dans les deux premiers chapitres nous développons des méthodes de bootstrap sur des estimateurs qui utilisent l'approche de "pré-moyennement". Il existe dans la littérature certaines applications du bootstrap aux données de hautes fréquences mais aucune n'est robuste à la présence d'erreurs de microstructure. Récemment Gonçalves et Meddahi (2009) ont montré la validité du wild bootstrap et du i.i.d. bootstrap pour la volatilité réalisée, mais dans un contexte où les prix des actifs sont sans frictions. Le premier chapitre de cette thèse a le mérite d'introduire dans cette littérature la première méthode de bootstrap robuste à la présence d'effets de microstructure. L'avantage du bootstrap est de permettre d'obtenir une inférence plus précise que celle avec la théorie asymptotique classique. Ainsi, nous nous sommes intéressés à l'estimateur de la volatilité intégrée proposé par Podolskij et Vetter (2009) utilisant des rendements "pré-moyennés". En effet pour cet estimateur les rendements "pré-moyennés" sont construits sur des blocs de rendements à hautes fréquences consécutifs qui ne se chevauchent pas. Le non chevauchement des blocs fait que les rendements "pré-moyennés" sont asymptotiquement indépendants, mais possiblement hétéroscédastiques. Ce qui motive l'application du wild bootstrap dans ce contexte. Nous montrons la validité théorique du bootstrap pour construire des intervalles de type percentile et percentile-t. A l'aide des simulations Monte Carlo nous montrons que le bootstrap améliore les propriétés en échantillon fini de l'estimateur de la volatilité intégrée par rapport aux résultats asymptotiques, pourvu que le choix de la variable externe soit fait de façon appropriée. Nous illustrons cette méthode en utilisant des données financières réelles.

Comme le premier chapitre de la thèse, le second chapitre s'intéresse à l'approche de "pré-moyennement". Mais contrairement au premier il se focalise sur l'estimateur de la volatilite intégrée proposé par Jacod et al. (2009). En effet cet estimateur a l'avantage d'être plus efficace comparativement à celui proposé par Podolskij et Vetter (2009). Pour cet estimateur les rendements "pré-moyennés" sont construits sur des blocs de rendements à haute fréquences consécutifs qui se chevauchent. Le chevauchement des blocs induit une forte dépendance dans la structure des rendements "pré-moyennés". Nous montrons que les rendements "pré-moyennés" sont m-dépendent avec m qui croît en fonction de la taille d'échantillon. Ceci motive l'application d'un bootstrap par bloc spécifique. Nous montrons qu'une application naïve du bootstrap par bloc suggéré par Bühlmann et Künsch (1995) n'est valide que lorsque la volatilité est constante. Nous argumentons que cela est dû à l'hétérogénéité dans la moyenne des rendements "pré-moyennés" au carré lorsque la volatilité est stochastique. Nous proposons donc une nouvelle procédure de bootstrap qui combine le wild bootstrap et le bootstrap par bloc, de telle sorte que la dépendance sérielle des rendements "pré-moyennés" est préservée à l'intérieur des blocs et la condition d'homogénéité nécessaire pour la validité du bootstrap est respectée. Sous des conditions de taille de bloc, nous montrons que cette méthode est convergente. Les simulations Monte Carlo montrent que le bootstrap améliore les propriétés en échantillon fini de l'estimateur de la volatilité intégrée par rapport aux résultats asymptotiques. Nous appliquons cette méthode sur des données financières réelles.

Contrairement aux deux premiers chapitres, le troisième ne s'intéresse pas à l'approche de "pré-moyennement". Il s'intéresse toujours aux données à haute fréquence, et à la volatilité réalisée mais dans un contexte différent. Généralement pour faire l'inférence statistique à l'aide des données financières de haute fréquence, les économètres ont l'habitude de supposer que la volatilité des prix des actifs est localement constante et utilisent certaines propriétés de la loi normale sur ces données. Nous étudions donc dans le troisième chapitre, dans quelle mesure et de quelle façon, l'approximation des distributions des estimateurs de mesures de co-volatilité sous l'hypothèse de Gaussianité locale des rendements peut être réalisée. Nous proposons une nouvelle méthode de bootstrap sous ces hypothèses. En particulier nous nous sommes intéressés à la volatilité réalisée et au beta réalisé. Nous montrons que la nouvelle méthode de bootstrap appliquée au beta réalisé était capable de répliquer les cummulants au deuxième ordre, tandis qu'il procurait un raffinement au troisième degré lorsqu'elle est appliquée à la volatilite réalisée. Ces résultats améliorent donc les résultats existants dans cette littérature, notamment ceux de Gonçalves et Meddahi (2009) et de Dovonon, Gonçalves et Meddahi (2013). Les simulations Monte Carlo montrent que le bootstrap améliore les propriétés en échantillon fini de l'estimateur de la volatilité intégrée par rapport aux résultats asymptotiques et les résultats de bootstrap existants. Enfin, nous illustrons toutes les méthodes développées dans cette thèse en utilisant des données financières réelles.

Chapter 1

Bootstrap inference for pre-averaged realized volatility based on non-overlapping returns

1.1 Introduction

The increasing availability of financial return series measured over higher and higher frequencies (e.g. every minute or every second) has revolutionized the field of financial econometrics over the last decade. Researchers and practitioners alike now routinely rely on high frequency data to estimate volatility (and functionals of it, such as regression and correlation coefficients).

One earlier popular estimator was realized volatility, computed as the sum of squared intraday returns. This is a consistent estimator of integrated volatility (a measure of the ex-post variation of asset prices over a given day) under quite general assumptions on the volatility process. However, one important assumption underlying the consistency of realized volatility is the assumption that markets are frictionless (so that asset prices are observed without any error). This assumption does not hold in practice. As the sampling frequency increases, market microstructure effects such as the existence of bid-ask bounds, rounding errors, discrete trading prices, etc, contribute to a discrepancy between the true efficient price process and the price observed by the econometrician (known as the market microstructure noise).

The negative impact of market microstructure effects on realized volatility is now an accepted fact in the econometrics literature of high frequency data. A number of alternative estimators have been proposed that take into account these effects (see e.g. Zhou (1996), Zhang et al. (2005), Hansen and Lunde (2006), Bandi and Russell (2008), Barndorff-Nielsen et al. (2008), Podolskij and Vetter (2009) and Jacod et al. (2009)). Although these estimators rely on a large number of high frequency returns, finite sample distortions associated with the first order normal approximation may persist even at large sample sizes, as shown by our simulations.

In this chapter, we consider the bootstrap as an alternative method of inference. We focus on the pre-averaging approach of Podolskij and Vetter (2009), where we first “average” the observed noisy returns over given blocks of non-overlapping observations, and then apply the standard realized volatility estimator to the pre-averaged returns. By averaging returns, the impact of the market microstructure noise is lessened, thus justifying realized volatility-like estimation on the pre-averaged returns. The class of statistics that we consider can be written (up to a bias term) as the (scaled) sum of squared pre-averaged returns (using an appropriate weighting function) computed over non-overlapping intervals. Our proposal is to bootstrap the pre-averaged returns.

Jacod et al. (2009) propose a generalization of the pre-averaging approach of Podolskij and Vetter (2009) which entails the use of overlapping intervals and the use of a more general weighting function for the pre-averaging of returns over these intervals. In this chapter, we consider the case of non-overlapping returns only. The main reason is that the structure of dependence of the pre-averaged returns is much simpler in this case as compared to the overlapping case, which simplifies inference significantly. In the non-overlapping case, the pre-averaged returns are independent asymptotically (as the number of blocks increases) but possibly heteroskedastic (due to stochastic volatility). Thus a wild bootstrap applied to the pre-averaged returns is asymptotically valid. In contrast, overlapping pre-averaged returns (as in Jacod et al. (2009)) are very strongly dependent because they rely on common returns. Therefore, the wild bootstrap is not appropriate and more sophisticated bootstrap methods are required. In particular, in Chapter 2, we show that a combination of the wild bootstrap with the blocks of blocks bootstrap of Bühlmann and Künsch (1995) (see also Künsch (1989), Politis and Romano (1992)) is asymptotically valid when applied to the pre-averaging estimator of Jacod et al. (2009). Although more generally applicable, the wild blocks of blocks bootstrap has the disadvantage of requiring the choice of a block size (in addition to the choice of the external random variable). For this reason, here we focus on the simpler non-overlapping case.

Our main contribution is to provide a proof of the validity of the wild bootstrap. Specifically, we follow the literature and model the observed price process as the sum of the true but latent price process (defined as a Brownian semimartingale process subject to stochastic volatility of a general nonparametric form) plus a noise term which captures the market microstructure noise. As in Podolskij and Vetter (2009), the noise is assumed i.i.d. Under these assumptions, the pre-averaged returns are asymptotically independent and play the role of the original returns in the realized volatility estimator when no market microstructure noise exists. Therefore, the proof of the validity of the wild bootstrap in the present context where market microstructure effects exist parallels the proof of the validity of the wild bootstrap in the context of Gonçalves and Meddahi (2009), where the wild bootstrap was proposed for realized volatility under no market microstructure effects. Nevertheless, an important difference between these two applications is the fact that the pre-averaging estimator of integrated volatility entails an analytical bias correction term. As it turns out, this bias correction is only important for the proper centering of the confidence intervals and does not impact the variance of the estimator. As a consequence, we show that no bias correction term is needed in the bootstrap world (because we can always center the bootstrap statistic at its own theoretical mean, without affecting the bootstrap variance). This simplifies the application of the bootstrap in this context and justifies an approach solely based on bootstrapping the pre-averaged returns (as the bias term typically depends on the highest available frequency returns, which we are not resampling in the proposed approach).

We first discuss conditions under which the wild bootstrap variance is a consistent estimator of the (conditional) variance of the pre-averaged realized volatility. Specifically, we show that a necessary condition for the consistency of the wild bootstrap variance is that $\mu_4^* - (\mu_2^*)^2 = \frac{2}{3}$, where $\mu_q^* \equiv E\,|v_j|^q$ and v_j denotes the external random variable used to generate the wild bootstrap pre-averaged returns $\bar{Y}_j^* = \bar{Y}_j \cdot v_j$, where $\bar{Y}_j$ are the pre-averaged returns. Under this condition, the bootstrap distribution of the scaled difference between the bootstrap pre-averaged realized volatility and its conditional mean is consistent for the (conditional) distribution of the pre-averaged realized volatility estimator. This result justifies the asymptotic validity of bootstrap percentile intervals for integrated volatility. Although this type of intervals does not promise asymptotic refinements over the first-order asymptotic approximation, they are easier to implement as they do not require an explicit estimator of the variance[1]. We then discuss the first-order asymptotic validity of bootstrap percentile-t intervals. In this case, we propose a consistent bootstrap variance estimator and show that the studentized bootstrap statistic based on this estimator is asymptotically normal for any choice of the external random variable, provided we center and scale the bootstrap statistic appropriately.

[1]In the univariate context considered here, the estimator of the variance of the pre-averaged realized volatility estimator is rather simple (it is given by a (scaled) version of the realized quarticity of pre-averaged returns), but this is not necessarily the case for other applications. For instance, for realized regression and realized correlation coefficients defined by the pre-averaging approach, the variance estimator is obtained by the delta method (whose finite sample properties are often poor) and the bootstrap percentile method coud be useful in that context.

We provide a set of Monte Carlo experiments that compare the finite sample performance of the bootstrap with the existing mixed normal approximation. Our results show that the choice of the external random variable is rather important in finite samples. In particular, percentile intervals that do not satisfy the moment condition $\mu_4^* - (\mu_2^*)^2 = \frac{2}{3}$ behave quite poorly in finite samples, confirming our theoretical result. In contrast, asymptotically valid percentile intervals behave similarly to the asymptotic theory-based intervals and both are dominated by percentile-t bootstrap intervals. Although percentile-t intervals are asymptotically valid for any choice of the external random variable, their finite sample performance is also influenced by this choice. Our results show that matching the first four cumulants (including the variance but also the mean, the skewness and the kurtosis) of the studentized statistic is important for good coverage properties. The optimal choice proposed by Gonçalves and Meddahi (2009) fails to do so when the sample size is small and therefore does not work well in the simulations. This suggests that a different choice may be optimal in the present context. Deriving such a choice would require the development of an Edgeworth expansion for the studentized statistic based on the pre-averaged realized volatility estimator and is outside the scope of this chapter. This is a non-trivial exercise given that the presence of the bias correction in the pre-averaged realized volatility estimator has an impact on the higher order cumulants, as our simulations shows. Instead, we show by simulation that a specific choice of the external random variable that does well in mimicking the first four cumulants of the statistic of interest has good finite sample coverage properties in the context of our Monte Carlo design.

The remainder of this chapter is organized as follows. In Section 1.2, we introduce the basic model and the main assumptions. Furthermore, we review the existing first-order asymptotic theory. We also introduce the Monte Carlo design underlying all simulations in the chapter and discuss the coverage probability results for the first-order asymptotic approach for nominal 95% two-sided symmetric intervals. In Section 1.3, we introduce our resampling method and prove its first-order asymptotic validity. In Section 1.4 we discuss the Monte Carlo results for bootstrap two-sided intervals. Section 1.5 contains an empirical application and Section 1.6 concludes. In the Appendix we give some technical results and present tables that illustrate the finite sample properties of the proposed procedures.

1.2 Setup, assumptions and review of existing results

1.2.1 Setup and assumptions

Let X denote the unobservable efficient log-price process defined on a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$. We model X as a Brownian semimartingale process defined by the equation

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s,\ t \geq 0, \tag{1.1}$$

where μ is a predictable locally bounded drift term, σ is an adapted càdlàg spot volatility process and W a standard Brownian motion. The object of interest is the quadratic variation of X given by

$$\int_0^T \sigma_s^2 ds,$$

also known as the integrated volatility. Without loss of generality, we let $T = 1$ and define $IV \equiv \int_0^1 \sigma_s^2 ds$ as the integrated volatility of X over a given time interval $[0, 1]$, which we think of as a given day.

The presence of market frictions such as price discreteness, rounding errors, bid-ask spreads, gradual response of prices to block trades, etc, prevent us from observing the true efficient price process X. Instead, we observe a noisy price process Y, given by

$$Y_t = X_t + \epsilon_t,$$

where ϵ_t represents the noise term that collects all the market microstructure effects. We assume that ϵ_t is i.i.d. and that ϵ_t is independent of X_t. Assumption 1 below collects these assumptions.

Assumption 1

(i) The noise component ϵ_t is i.i.d. $(0, \omega^2)$ with $E\left|\epsilon_t\right|^{8+\varepsilon} < \infty$ for some $\varepsilon > 0$.

(ii) ϵ_t is independent from the latent log-price X_t.

Assumption 1 is standard in the literature on market microstructure noise robust estimators of integrated volatility (see, among others, Zhang et al. (2005), Barndorff-Nielsen et al. (2008), Podolskij and Vetter (2009)). Nevertheless, empirically the i.i.d. assumption on ϵ and the independence between X and ϵ may be too strong a set of assumptions, especially at the highest frequencies. See e.g. Hansen and Lunde (2006), Zhang et al. (2011b), Diebold and Strasser (2012) for more on this issue. For simplicity, we will maintain these assumptions throughout.

Although for consistency of the pre-averaging estimator, $4+\varepsilon$ moments of ϵ_t suffice (see, in particular, Theorem 1 of Podolskij and Vetter (2009) with $r=2$ and 0), here we impose a stronger moment condition that requires the existence of $8+\varepsilon$ moments. This is because we are interested in approximating the entire distribution of the studentized statistic based on the pre-averaging realized volatility estimator and we need a consistent estimator of its conditional variance. Consistency of the variance estimator requires this strenghtening of the moment condition (see again Theorem 1 of Podolskij and Vetter (2009) with $r=4$ and $l=0$). Note that in contrast to Podolskij and Vetter (2009), we do not need to impose a Gaussianity assumption on ϵ, nor do we need to restrict the volatility process σ to be a Brownnian semi-martingale. These assumptions are needed when studying the asymptotic properties of bipower or multipower pre-averaging statistics but can be dispensed with in the case of squared averaged returns (see Vetter (2008), p.49, for more details on this).

1.2.2 The pre-averaging approach

Suppose we observe Y at regular time points $\frac{i}{n}$, for $i=0,\ldots,n$, from which we compute n intraday returns at frequency $\frac{1}{n}$,

$$r_i \equiv Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}, \quad i=1,\ldots,n.$$

Given that $Y = X + \epsilon$, we can write

$$r_i = \left(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\right) + \left(\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}\right) \equiv r_i^e + \Delta\epsilon_i,$$

where $r_i^e = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ denotes the $\frac{1}{n}$-frequency return on the efficient price process.

We can show that

$$r_i = r_i^e + \Delta\epsilon_i = O_P\left(\frac{1}{\sqrt{n}}\right) + O_P(1). \tag{1.2}$$

Since X follows a stochastic volatility model given by (1.1), r_i^e is (conditionally on the path of σ and μ) independent and heteroskedastic with (conditional) variance given by $\int_{(i-1)/n}^{i/n} \sigma_s^2 ds$. The order of magnitude of r_i^e is thus $O_P\left(\frac{1}{\sqrt{n}}\right)$. In contrast, under Assumption 1, the difference $\Delta\epsilon_i \equiv \epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}$ is an $MA(1)$ process whose order of magnitude is $O_P(1)$.

The decomposition in (1.2) shows that the noise completely dominates the observed return process as $n\to\infty$. This in turn implies that the usual realized volatility estimator is biased and inconsistent.

Moreover, even though the efficient returns r_i^e are conditionally independent, this is no longer the case for the observed returns. More specifically, the i.i.d. assumption on ϵ_t implies that the observed returns r_i are (conditionally) one-dependent due to the $MA\,(1)$ structure induced by the i.i.d. noise process.

Several approaches have been considered in the literature. Zhang et al. (2005) proposed a subsampling approach and derived the two times scale realized volatility estimator. This estimator amounts to using a linear combination of realized volatility estimators computed on subsamples (the slow scale) and an analytical bias correction term that relies on a realized volatility computed on a fast scale. Barndorff-Nielsen et al. (2008) proposed the realized kernel estimators, where linear combinations of autocovariances are considered. More recently, Podolskij and Vetter (2009) introduced the pre-averaging approach based on non-overlapping blocks. This was further generalized in Jacod et al. (2009) to allow for overlapping blocks.

In this chapter we focus on bootstrapping the pre-averaged realized volatility estimator of Podolskij and Vetter (2009). As we mentioned before, our proposal is to bootstrap the pre-averaged returns. By focusing only on non-overlapping intervals, we can apply the wild bootstrap method to the pre-averaged returns. The dependence structure of the pre-averaged returns becomes much stronger under overlapping intervals and invalidates the use of the wild bootstrap. See Chapter 2 for a bootstrap method that is valid in this context and which combines the wild bootstrap with a blocks of blocks bootstrap.

Next we describe the pre-averaging approach of Podolskij and Vetter (2009). This approach depends on two tuning parameters K and L, which denote two different block sizes. Specifically, let K denote the size of a block of K consecutive $\frac{1}{n}$-horizon returns. Within each non-overlapping block of size K, we consider the set of all overlapping blocks of size L, where L is a fraction of K. For a given (non-overlapping) block of size K, there will be such $K - L + 1$ blocks of size L.

Assume that n/K is an integer so that the number of non-overlapping blocks of size K is n/K. For $j = 1, \ldots, n/K$, the pre-averaged return $\bar{Y}_j$ is obtained as follows:

$$\bar{Y}_j = \frac{1}{K - L + 1} \sum_{i=(j-1)K}^{jK-L} \left(\sum_{l=1}^{L} r_{l+i} \right).$$

This amounts to computing the sum of $\frac{1}{n}$-horizon returns over each block of size L and then averaging the result over all possible such overlapping blocks. An alternative expression for $\bar{Y}_j$ is as follows:

$$\bar{Y}_j = \sum_{i=1}^{K} g\,(i, K, L)\, r_{i+(j-1)K},$$

where for every $i = 1, \dots, K$, the weighting function $g(i, K, L)$ is defined as

$$g(i, K, L) = \begin{cases} \frac{i}{K-L+1}, & \text{if } \ i \in \{1, \dots, L\} \\ \frac{L}{K-L+1}, & \text{if } \ i \in \{L+1, \dots, K-L\} \\ \frac{K-i+1}{K-L+1}, & \text{if } \ i \in \{K-L+1, \dots, K\} \end{cases},$$

and where we can show that $\sum_{i=1}^{K} g(i, K, L) = L$.

The effect of pre-averaging is to reduce the impact of the noise in the pre-averaged return. Specifically, we can show that by pre-averaging returns over blocks of size K in this particular manner, we reduce the variance by a factor of about $\frac{1}{K}$. To be more precise, Podolskij and Vetter (2009) show that

$$\bar{Y}_j = \bar{r}_j^e + \Delta\bar{\epsilon}_j = O_P\left(\sqrt{\frac{L}{n}}\right) + O_P\left(\frac{1}{\sqrt{K-L}}\right), \tag{1.3}$$

where $\bar{r}_j^e$ and $\Delta\bar{\epsilon}_j$ denote the pre-averaged versions of the efficient returns and the difference of the noise process, respectively. Thus, comparing (1.2) with (1.3), we see that pre-averaging manages to reduce the impact of the noise from $O_P(1)$ to $O_P\left(\frac{1}{\sqrt{K-L}}\right)$. Since L is a fraction of K, i.e. $L \sim \frac{1}{c_2}K$, for some $c_2 > 1$, the order of magnitude of the noise in (1.3) is $O_P\left(\frac{1}{\sqrt{K}}\right)$. The overall implication is that we can compute a realized volatility-like estimator on the pre-averaged returns $\bar{Y}_j$. This is the essence of the pre-averaging approach.

To give the explicit formula of the pre-averaging realized volatility estimator of Podolskij and Vetter (2009), we need to introduce some additional notation. In particular, we let

$$L = \frac{1}{c_2}K, \tag{1.4}$$

with $c_2 > 1$, and

$$K = c_1 c_2 \sqrt{n}, \tag{1.5}$$

where $c_1 > 0$, and c_1 and c_2 are two tuning parameters that need to be chosen. These choices of K and L imply that the two terms in (1.3) are balanced and equal to $O_P\left(n^{-1/4}\right)$.

Under Assumption 1, and assuming that K and L satisfy the conditions (1.4) and (1.5), respectively, Podolskij and Vetter (2009) [cf. Theorem 1] show that

$$p\lim_{n\to\infty}\left(\sum_{j=1}^{n/K} \bar{Y}_j^2\right) = \frac{\nu_1}{c_1 c_2}\int_0^1 \sigma_s^2 ds + \frac{\nu_2}{c_1 c_2}\omega^2,$$

where $\omega^2 = Var(\epsilon_i)$ and where

$$\nu_1 = \frac{c_1 \left(3c_2 - 4 + \max\left((2-c_2)^3, 0\right)\right)}{3(c_2-1)^2}, \ \nu_2 = \frac{2(\min((c_2-1), 1))}{c_1 (c_2-1)^2}.$$

Two implications can be obtained from this result. First, the particular weighting scheme induced by the pre-averaging approach introduces a scaling factor given by $\frac{\nu_1}{c_1 c_2}$ when estimating $\int_0^1 \sigma_s^2 ds$. This implies that we need to scale $\sum\limits_{j=1}^{n/K} \bar{Y}_j^2$ by $\frac{c_1 c_2}{\nu_1}$. Second, although the pre-averaging approach reduces the order of magnitude of the noise, it does not completely eliminate its influence. In particular,

$$p \lim_{n\to\infty} \left(\frac{c_1 c_2}{\nu_1} \sum_{j=1}^{n/K} \bar{Y}_j^2\right) = \int_0^1 \sigma_s^2 ds + \underbrace{\frac{\nu_2}{\nu_1}\omega^2}_{\text{Bias term}},$$

where the bias term is proportional to the variance of the noise ω^2. A consistent estimator of ω^2 is given by the realized volatility estimator computed on the n highest frequency returns r_i, divided by $2n$, i.e.

$$\hat{\omega}^2 = \frac{\sum_{i=1}^n r_i^2}{2n} \to^P \omega^2.$$

This suggests the following consistent estimator of integrated volatility:

$$PRV_n = \frac{c_1 c_2}{\nu_1} \underbrace{\sum_{j=1}^{n/K} \bar{Y}_j^2}_{RV\text{-like estimator}} - \underbrace{\frac{\nu_2}{\nu_1}\hat{\omega}^2}_{\text{bias correction term}}.$$

1.2.3 First-order asymptotic distribution theory

Under Assumption 1, and assuming that K and L are chosen according to (1.4) and (1.5), Podolskij and Vetter (2009) (cf. Corollary 1) show that

$$\frac{n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{V}} \to^{st} N(0,1), \tag{1.6}$$

where $\to^{st}$ denotes stable convergence (see Christensen and al. (2009), p. 119 for a definition of stable convergence), and

$$V = \frac{2c_1^2 c_2^2}{\nu_1^2} \int_0^1 \left(\nu_1 \sigma_s^2 + \nu_2 \omega^2\right)^2 ds$$

is the conditional variance of PRV_n.

By Theorem 1 of Podolskij and Vetter (2009), a consistent estimator of V is given by

$$\hat{V}_n = \frac{2c_1^2 c_2^2}{3\nu_1^2}\sqrt{n}\sum_{j=1}^{n/K}\left|\bar{Y}_j\right|^4.$$

This estimator has the form of a realized quarticity estimator applied to the pre-averaged returns $\bar{Y}_j$. Together with the CLT result (1.6), it implies that (cf. equation (3.19) in Podolskij and Vetter (2009))

$$T_n \equiv \frac{n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{\hat{V}_n}} \to^d N(0,1).$$

We can use this feasible asymptotic distribution result to build confidence intervals for integrated volatility. In particular, a two-sided feasible $100(1-\alpha)\%$ level interval for $\int_0^1 \sigma_s^2 ds$ is given by:

$$IC_{Feas,1-\alpha} = \left(PRV_n - z_{1-\alpha/2}n^{-1/4}\sqrt{\hat{V}_n}, PRV_n + z_{1-\alpha/2}n^{-1/4}\sqrt{\hat{V}_n}\right),$$

where $z_{1-\alpha/2}$ is such that $\Phi\left(z_{1-\alpha/2}\right) = 1-\alpha/2$, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. For instance, $z_{0.975} = 1.96$ when $\alpha = 0.05$.

1.2.4 Finite sample properties of the feasible asymptotic approach

In this section we assess by Monte Carlo simulation the accuracy of the feasible asymptotic theory of the pre-averaging approach of Podolskij and Vetter (2009). We find that this approach leads to important coverage probability distortions when returns are not sampled too frequently. This motivates the bootstrap as an alternative method of inference in this context.

We consider two data generating processes in our simulations. First, following Zhang et al. (2005), we use the one-factor stochastic volatility (SV1F) model of Heston (1993) as our data-generating process, i.e.

$$dX_t = (\mu - \nu_t/2)\,dt + \sigma_t dB_t,$$

and

$$d\nu_t = \kappa(\alpha - \nu_t)\,dt + \gamma(\nu_t)^{1/2}\,dW_t,$$

where $\nu_t = \sigma_t^2$, B and W are two Brownian motions, and we assume $Corr(B,W) = \rho$.

The parameter values are all annualized. In particular, we let $\mu = 0.05/252$, $\kappa = 5/252$, $\alpha = 0.04/252$, $\gamma = 0.05/252$, $\rho = -0.5$. The size of the market microstucture noise is an important parameter. We follow Barndorff-Nielsen et al. (2009) and model the noise magnitude as $\xi^2 = \omega^2/\sqrt{\int_0^1 \sigma_s^4 ds}$. We fix ξ^2 be equal to 0.0001, 0.001 and 0.01, and let $\omega^2 = \xi^2\sqrt{\int_0^1 \sigma_s^4 ds}$. These values are motivated by the empirical study of Hansen and Lunde (2006), who investigate 30 stocks of Dow Jones Industrial Average.

We also consider the two-factor stochastic volatility (SV2F) model analyzed by Barndorff-Nielsen et al. (2009), where [2]

$$
\begin{aligned}
dX_t &= \mu dt + \sigma_t dB_t, \\
\sigma_t &= s\text{-}exp\left(\beta_0 + \beta_1 \tau_{1t} + \beta_2 \tau_{2t}\right), \\
d\tau_{1t} &= \alpha_1 \tau_{1t} dt + dB_{1t}, \\
d\tau_{2t} &= \alpha_2 \tau_{2t} dt + \left(1 + \phi \tau_{2t}\right) dB_{2t}, \\
corr\left(dW_t, dB_{1t}\right) &= \varphi_1,\ corr\left(dW_t, dB_{2t}\right) = \varphi_2.
\end{aligned}
$$

We follow Huang and Tauchen (2005) and set $\mu = 0.03$, $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.00137$, $\alpha_2 = -1.386$, $\phi = 0.25$, $\varphi_1 = \varphi_2 = -0.3$. We initialize the two factors at the start of each interval by drawing the persistent factor from its unconditional distribution, $\tau_{10} \sim N\left(0, \frac{-1}{2\alpha_1}\right)$, by starting the stronlgly mean-reverting factor at zero.

We simulate data for the unit interval $[0, 1]$ and normalize one second to be 1/23400, so that $[0, 1]$ is thought to span 6.5 hours. The observed Y process is generated using an Euler scheme. We then construct the $\frac{1}{n}$-horizon returns $r_i \equiv Y_{i/n} - Y_{(i-1)/n}$ based on samples of size n.

The pre-averaging approach requires the choice of the tuning parameters c_1 and c_2. Podolskij and Vetter (2009) give the optimal values of c_1 and c_2 that minimize the conditional variance V of the PRV_n estimator when the volatility process is constant. In our simulations, we followed Podolskij and Vetter (2009) and let $c_2 = 1.6$ and $c_1 = 1$. These choices may not be optimal under stochastic volatility, but since we will compute the bootstrap statistics using these same values, they allow for a meaningful comparison of the different intervals for integrated volatility (asymptotic theory-based and bootstrap intervals).

Table 3.1 gives the actual rates of 95% confidence intervals of integrated volatility for the SV1F and the SV2F models, respectively, computed over 10,000 replications. Results are presented for eight different samples sizes: n = 23400, 11700, 7800, 4680,

[2]The function s-exp is the usual exponential function with a linear growth function splined in at high values of its argument: $s\text{-}exp(x) = exp(x)$ if $x \leq x_0$ and $s\text{-}exp(x) = \frac{exp(x_0)}{\sqrt{x_0 - x_0^2 + x^2}}$ if $x > x_o$, with $x_0 = log(1.5)$.

1560, 780, 390 and 195, corresponding to "1-second", "2-second", "3-second", "5-second", "15-second", "30-second", "1-minute" and "2-minute" frequencies (this table also includes results for the bootstrap methods but those results will be discussed later in Section 3.)

For the two models, all intervals tend to undercover. The degree of undercoverage is especially large for smaller values of n, when sampling is not too frequent. The SV2F model exhibits overall larger coverage distortions than the SV1F model, for all sample sizes. Results are not very sensitive to the noise magnitude.

1.3 The bootstrap

In this section we provide a bootstrap method for inference on integrated volatility based on the pre-averaging approach of Podoslkij and Vetter (2009). Our proposal is to bootstrap the pre-averaged returns $\bar{Y}_j$, $j = 1, \ldots, n/K$. Because non-overlapping intervals are used, the pre-averaged returns $\bar{Y}_j$ are asymptotically independent, as $n \to \infty$. In fact, we can show that they are one-dependent, i.e. $\bar{Y}_j$ is independent of $\bar{Y}_m$ whenever $|m - j| > 1$. Moreover, the amount of dependence between two consecutive squared pre-averaged returns is very small and it is only due to edge effects. Specifically, $Cov\left(\bar{Y}_j^2, \bar{Y}_{j+1}^2\right) = O\left(\frac{1}{n^2}\right) = o\left(1\right)$ as $n \to \infty$.

Since pre-averaged returns are asymptotically independent but possibly heteroskedastic (due to the fact that volatility is time-varying) a wild bootstrap approach is appropriate. The wild bootstrap method was introduced by Wu (1986), and further studied by Liu (1988) and Mammen (1993), in the context of cross-section linear regression models subject to unconditional heteroskedasticity in the error term. Gonçalves and Meddahi (2009) applied the wild bootstrap method in the context of realized volatility under no market microstructure noise. Our approach here follows Gonçalves and Meddahi (2009), but instead of bootstrapping the $\frac{1}{n}$-horizon raw returns r_i, we propose to bootstrap the pre-averaged returns $\bar{Y}_j$.

The bootstrap pseudo-data is given by

$$\bar{Y}_j^* = \bar{Y}_j \cdot v_j, \quad j = 1, \ldots, n/K,$$

where the external random variable v_j is an i.i.d. random variable independent of the data and whose moments are given by $\mu_q^* \equiv E^* \left|v_j\right|^q$. As usual in the bootstrap literature, P^* (E^* and Var^*) denotes the probability measure (expected value and variance) induced by the bootstrap resampling, conditional on a realization of the original time series. In addition, for a sequence of bootstrap statistics Z_n^*, we write $Z_n^* = o_{P^*}\left(1\right)$ in probability, or $Z_n^* \to^{P^*} 0$, as $n \to \infty$, in probability, if for any $\varepsilon > 0$, $\delta > 0$,

$\lim_{n\to\infty} P\left[P^*\left(|Z_n^*| > \delta\right) > \varepsilon\right] = 0$. Similarly, we write $Z_n^* = O_{P^*}(1)$ as $n \to \infty$, in probability if for all $\varepsilon > 0$ there exists a $M_\varepsilon < \infty$ such that $\lim_{n\to\infty} P\left[P^*\left(|Z_n^*| > M_\varepsilon\right) > \varepsilon\right] = 0$. Finally, we write $Z_n^* \to^{d_*} Z$ as $n \to \infty$, in probability, if conditional on the sample, Z_n^* weakly converges to Z under P^*, for all samples contained in a set with probability P converging to one.

The bootstrap pre-averaged realized volatility estimator is given by

$$PRV_n^* = \frac{c_1 c_2}{\nu_1} \sum_{j=1}^{n/K} \bar{Y}_j^{*2}.$$

Although the pre-averaged realized volatility estimator PRV_n contains a bias correction term, we do not consider bias correction in the bootstrap world. The reason is twofold. First, our goal is not to estimate consistently the integrated volatility using the bootstrap. Instead, our goal is to use the bootstrap to approximate the distribution of statistics based on PRV_n, for instance we would like to approximate the distribution of the t-statistic T_n defined in the previous section. We can easily show that

$$E^*\left(PRV_n^*\right) = \mu_2^* \frac{c_1 c_2}{\nu_1} \sum_{j=1}^{n/K} \bar{Y}_j^2.$$

This is a biased estimator of integrated volatility, but we can correctly center our bootstrap statistics using this theoretical bootstrap mean. Since the bias correction term does not affect the variance of the pre-averaging estimator, as long as the bootstrap method is able to consistently estimate this variance, no bias correction is needed in the bootstrap world. The second reason why we do not consider bootstrap bias correction is that the bootstrap bias correction term would involve the bootstrap highest frequency returns r_i^*, which are not available under our proposed method.

We can show that

$$Var^*\left(n^{1/4} PRV_n^*\right) = \left(\mu_4^* - \mu_2^{*2}\right) \underbrace{\frac{c_1^2 c_2^2}{\nu_1^2} \sqrt{n} \sum_{j=1}^{n/K} \left|\bar{Y}_j\right|^4}_{\equiv \frac{3}{2}\hat{V}_n}.$$

It follows then that a sufficient condition for the bootstrap to provide a consistent estimator of the conditional variance of $n^{1/4} PRV_n$ is that $\mu_4^* - \mu_2^{*2} = \frac{2}{3}$. Under this condition, the bootstrap can be used to approximate the quantiles of the distribution of the root

$$n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right),$$

thus justifying the construction of bootstrap percentile confidence intervals.

These results are summarized in the following theorem.

Theorem 1.3.1. *Suppose Assumption 1 holds and let K and L satisfy the conditions (1.4) and (1.5), respectively. Suppose that $\left\{\bar{Y}_j^* = \bar{Y}_j \cdot v_j : j = 1, \ldots, n/K\right\}$, where $v_j \sim$ i.i.d. such that for any $\delta > 0$, $\mu_{2(2+\delta)}^* = E^* |v_j|^{2(2+\delta)} < \infty$. If $\mu_4^* - (\mu_2^*)^2 = \frac{2}{3}$, then as $n \to \infty$,*

(1) $V_n^* \equiv Var^* \left(n^{1/4} PRV_n^*\right) \xrightarrow{P} V \equiv \frac{2c_1^2 c_2^2}{\nu_1^2} \int_0^1 \left(\nu_1 \sigma_s^2 + \nu_2 \omega^2\right)^2 ds.$

(2) $\sup_{x \in \mathbb{R}} \left| P^* \left(n^{1/4} \left(PRV_n^* - E^* \left(PRV_n^*\right)\right) \leq x\right) - P \left(n^{1/4} \left(PRV_n - \int_0^1 \sigma_s^2 ds\right)\right) \right| \xrightarrow{P} 0.$

An example of a random variable that satisfies the condition $\mu_4^* - \mu_2^{*2} = \frac{2}{3}$ is

$$v_j \sim \text{i.i.d. } N\left(0, \sqrt{3}/3\right).$$

Theorem 1.3.1 justifies using the wild bootstrap to construct bootstrap percentile intervals for integrated volatility. Specifically, a $100\,(1-\alpha)\,\%$ symmetric bootstrap percentile interval for integrated volatility based on the bootstrap is given by

$$IC_{perc,1-\alpha}^* = \left(PRV_n - n^{-1/4} p_{1-\alpha}^*, PRV_n + n^{-1/4} p_{1-\alpha}^*\right), \tag{1.7}$$

where $p_{1-\alpha}^*$ is the $1-\alpha$ quantile of the bootstrap distribution of $\left|n^{1/4}\left(PRV_n^* - E^*\left(PRV_n^*\right)\right)\right|$.

Bootstrap percentile intervals do not promise asymptotic refinements. Next, we propose a consistent bootstrap variance estimator that allows us to form bootstrap percentile-t intervals. More specifically, we can show that the following bootstrap variance estimator consistently estimates V_n^* for any choice of the external random variable v_j:

$$\hat{V}_n^* = \frac{\mu_4^* - \mu_2^{*2}}{\mu_4^*} \frac{c_1^2 c_2^2}{\nu_1^2} n^{1/2} \sum_{j=1}^{n/K} \bar{Y}_j^{*4}.$$

Our proposal is to use this estimator to construct a bootstrap studentized statistic,

$$T_n^* \equiv \frac{n^{1/4}\left(PRV_n^* - E^*\left(PRV_n^*\right)\right)}{\sqrt{\hat{V}_n^*}},$$

the bootstrap analogue of T_n.

Theorem 1.3.2. *Suppose Assumption 1 holds such that for any $\delta > 0$, $E|\epsilon_t|^{2(8+\delta)} < \infty$, and let K and L satisfy the conditions (1.4) and (1.5), respectively. Suppose that $\left\{\bar{Y}_j^* = \bar{Y}_j \cdot v_j : j = 1, \ldots, n/K\right\}$, where $v_j \sim$ i.i.d. such that $\mu_8^* = E^* |v_j|^8 < \infty$. It follows that as $n \to \infty$, $\sup_{x \in \mathbb{R}} \left|P^*\left(T_n^* \leq x\right) - P\left(T_n \leq x\right)\right| \xrightarrow{P} 0$.*

Theorem 1.3.2 justifies constructing bootstrap percentile-t intervals. In particular, a $100\,(1-\alpha)\,\%$ symmetric bootstrap percentile-t interval for integrated volatility is given by

$$IC^{*}_{perc-t,1-\alpha} = \left(PRV_n - q^{*}_{1-\alpha} n^{-1/4}\sqrt{\hat{V}_n}, PRV_n + q^{*}_{1-\alpha} n^{-1/4}\sqrt{\hat{V}_n}\right), \qquad (1.8)$$

where $q^{*}_{1-\alpha}$ is the $(1-\alpha)$-quantile of the bootstrap distribution of $|T^{*}_n|$. The first order asymptotic validity of the bootstrap requires a strenthening of the moment condition on ϵ_t when applied to the feasible statistic T_n.

1.4 Monte Carlo results for the bootstrap

In this section, we compare the finite sample performance of the bootstrap with the first order asymptotic theory for confidence intervals of integrated volatility. In our simulations, bootstrap intervals use 999 bootstrap replications for each of the 10,000 Monte Carlo replications.

We consider bootstrap percentile and bootstrap percentile-t intervals, computed at the 95% level using (1.7) and (1.8), respectively.

To generate the bootstrap data we use three different external random variables.

WB1 $v_j \sim$ i.i.d. $N\left(0, \sqrt{3}/3\right)$, implying that $\mu^{*}_2 = \sqrt{3}/3$ and $\mu^{*}_4 = 1$.

WB2 A two point distribution $v_j \sim$ i.i.d. such that:

$$v_j = \begin{cases} \left(\frac{2}{3}\right)^{1/4} \frac{-1+\sqrt{5}}{2}, & \text{with} \quad \text{prob} \quad p = \frac{\sqrt{5}-1}{2\sqrt{5}} \\ \left(\frac{2}{3}\right)^{1/4} \frac{-1-\sqrt{5}}{2}, & \text{with} \quad \text{prob} \quad 1-p = \frac{\sqrt{5}+1}{2\sqrt{5}} \end{cases},$$

for which $\mu^{*}_2 = 2\sqrt{2/3}$ and $\mu^{*}_4 = 10/3$.

WB3 The two point distribution proposed by Gonçalves and Meddahi (2009), where $v_j \sim$ i.i.d. such that:

$$v_j = \begin{cases} \frac{1}{5}\sqrt{31+\sqrt{186}}, & \text{with} \quad \text{prob} \quad p = \frac{1}{2} - \frac{3}{\sqrt{186}} \\ -\frac{1}{5}\sqrt{31-\sqrt{186}}, & \text{with} \quad \text{prob} \quad 1-p \end{cases},$$

for which we have $\mu^{*}_2 = 1$ and $\mu^{*}_4 = 31/25$.

The condition $\mu^{*}_4 - \left(\mu^{*}_2\right)^2 = \frac{2}{3}$ is satisfied for the first two choices (WB1 and WB2) but not for WB3. The implication is that WB1 and WB2 are valid for percentile intervals but not WB3. Note however that all three choices of v_j are asymptotically valid when used to construct bootstrap percentile-t intervals.

Table 3.1 shows the actual coverage probability rates of nominal 95% symmetric bootstrap intervals for integrated volatility based on WB1, WB2 and WB3 for each of the two models (SV1F and SV2F). Both percentile and percentile-t intervals are considered. Results based on the asymptotic normal distribution are also included (under the label CLT). As already discussed in Section 2.4, results are not very sensitive to the choice of ξ^2 and distortions are larger (both based on asymptotic theory and on the bootstrap) for the SV2F than for the SV1F model. These trends are also present for the bootstrap.

Starting with the bootstrap percentile intervals, we see that these are close to the CLT-based intervals for WB1 and WB2 (when the condition $\mu_4^* - (\mu_2^*)^2 = \frac{2}{3}$ is satisfied) whereas coverage rates for percentile intervals based on WB3 are systematically much lower than 95% even for the largest sample sizes. This confirms the theoretical prediction of asymptotic invalidity for these intervals. The results also confirm that the bootstrap percentile intervals do not outperform the asymptotic theory-based intervals. Nevertheless, choosing v_j to match the variance of the pre-averaging estimator may result in better percentile-t intervals, as a comparison the different bootstrap methods shows for this type of intervals. Specifically, although WB2 and WB3 both undercover for smaller sample sizes, WB2 outperforms WB3 significantly for the smaller samples sizes. For instance, for SV1F, WB3 covers IV 81.41% of the time when $n = 195$ whereas WB2 does so 91.05%. These rates decrease to 71.89% and 86.78% for the SV2F model, respectively. In contrast, the WB1 method covers IV with a rate equal to 97.91% for SV1F and 94.72% for SV2F, when $n = 195$. In general, the results show that percentile-t intervals based on WB1 are too conservative, yielding coverage rates larger than 95%, especially for the SV1F model. WB2 intervals tend to be closer to the desired nominal level than the WB3 method, without being conservative. Overall, the results suggest that the choice of v_j is important in finite samples.

In order to gain further insight into why the different choices of v_j matter in finite samples, we computed the first four cumulants of T_n and of its bootstrap analogue T_n^*. The results are presented in Table 3.2, which also reports the coverage rates of symmetric intervals based on these studentized statistics. Results are only given for $\xi^2 = 0.01$. For T_n, we report the mean, the standard error, the excess skewness and the excess kurtosis across the 10,000 simulations. For T_n^*, the numbers correspond to the average value (across the 10,000 simulations) of the bootstrap mean, standard error, excess skewness and excess kurtosis computed for each simulation across the 999 bootstrap replications.

Starting with T_n, the results show that this statistic is centered at a negative value across the different sample sizes. The negative bias decreases as n increases, but it can be quite large when n is small. Since the asymptotic normal distribution is centered at zero, it completely misses this downward bias. We can also see that the finite sample distribution of T_n is more dispersed than the $N(0,1)$ distribution (its standard error is larger than 1), and that it is strongly negatively skewed (the excess skewness is very negative) and fat-tailed (the excess kurtosis is positive). All these features explain the undercoverage of the CLT approach. In contrast, the bootstrap cumulants of T_n^* replicate to a better degree the finite sample patterns of the four cumulants of T_n depending on the choice of v_j. Specifically, we can see that the three choices of v_j typically induce a negative bias as well as negative excess skewness and positive excess kurtosis (an exception is WB3 for the smaller sample sizes). Nevertheless, WB1 implies too strong a correction. For instance, the bias of T_n^* is more negative than it should be on average as well as its excess skewness. This means that the bootstrap distribution of T_n^* is on average to the left of the finite sample distribution of T_n, resulting in too large a critical value, which explains the overcoverage problem noted in Table 3.1. In contrast, for the smaller sample sizes, WB2 and WB3 imply too little a correction in terms of the bias, which implies that these bootstrap distributions are on average centered to the right of the true distribution of T_n. This contributes to too small a critical value and to some undercoverage.

Overall, the results suggest that WB3 does a poorer job at capturing the first four cumulants than WB2, especially for the smaller sample sizes. This suggests that the optimal choice of v_j proposed by Gonçalves and Meddahi (2009) in the context of realized volatility without market microstructure noise is no longer optimal in the context of pre-averaging realized volatility. The presence of the bias correction term in the definition of PRV_n implies that the Edgeworth expansions derived in Gonçalves and Meddahi (2009) do not apply in the pre-averaging approach considered here. Thus, although bias correction does not have an impact to first order on the asymptotic variance of PRV_n, it likely has an impact on the higher order cumulants, as our Monte Carlo simulation results suggest. Deriving the optimal choice of the external random variable in this context is an interesting research question which we will consider elsewhere.

1.5 Empirical results

As a brief illustration, in this section we implement the proposed wild bootstrap method to real high frequency data, and compare it to the existing feasible asymptotic procedure of Podolskij and Vetter (2009). The data consists of transaction log prices of General Electric (GE) shares carried out on the New York Stock Exchange (NYSE) in December 2011. For each day, we consider data from the regular exchange opening hours from time stamped between 9:30 a.m. until 4 p.m. Our procedure for cleaning the data is exactly identical to that used by Barndorff-Nielsen et al. (2008).

We implement the pre-averaged realized volatility estimator of Jacod et al. (2009)

on returns recorded every S transactions, where S is selected each day so that there are approximately 1493 observations a day. This means that on average these returns are recorded roughly every 15 seconds. Table 3.3 in the Appendix provides the number of transactions per day and the sample size for the pre-averaged returns. The pre-averaged realized volatility estimator is implemented with $c_2 = 1.6$ and $c_1 = 1$.

We consider bootstrap percentile-t intervals, computed at the 95% level using (1.8), where v_j is generated using WB2 (our best choice according to the Monte Carlo simulations). The results are displayed in Figure 3.1 in terms of daily 95% confidence intervals (CIs) for integrated volatility. Two types of intervals are presented: our proposed wild bootstrap method and the existing feasible asymptotic procedure Podolskij and Vetter (2009). The pre-averaged realized volatility estimate is in the center of both confidence intervals by construction.

The confidence intervals for IV based on the bootstrap method are usually wider than the confidence intervals using the feasible asymptotic theory. Nevertheless, as our Monte Carlo simulations showed, the latter typically have undercoverage problems whereas the bootstrap intervals have coverage rates closer to the desired level. Therefore if the goal is to control the coverage probability, shorter intervals are not necessarily better. The figures also show a lot of variability in the daily estimate of integrated volatility.

1.6 Conclusion

In this chapter we propose the wild bootstrap as a method of inference for integrated volatility in the context of the pre-averaged realized volatility estimator proposed by Podolskij and Vetter (2009). The wild bootstrap is motivated by the fact that non-overlapped pre-averaged returns are asymptotically independent but possibly heteroskedastic (in the context of stochastic volatility models). We provide a set of conditions under which this method is asymptotically valid to first order. Both percentile and percentile-t bootstrap intervals are considered. Our Monte Carlo simulations show that the bootstrap can improve upon the mixed Gaussian inference derived by Podolskij and Vetter (2009) provided we choose the external random variable appropriately.

An important question for future research is the optimal choice of the external random variable in this context. This is not an easy question because it requires developing Edgeworth expansions for the statistics of interest in the original sample and the bootstrap samples. Since the pre-averaged realized volatility estimator depends on a bias correction term, its Edgeworth expansion will reflect the contribution of this term at higher orders and render the analysis rather complex. We plan on investigating this issue in future work.

Chapter 2

Bootstrapping pre-averaged realized volatility under market microstructure noise

2.1 Introduction

Estimation of integrated volatility is complicated by the existence of market microstructure noise. This noise represents the discrepancy between the true efficient price of an asset and its observed counterpart and is caused by a multitude of market microstructure effects (such as bid-ask bounds, the discreteness of price changes and the existence of rounding errors, the gradual response of prices to a block trade, the existence of data recording errors such as prices entered as zero, misplaced decimal points, etc).

Realized volatility, computed as the sum of squared intraday returns, is not consistent for integrated volatility under the presence of market microstructure noise. This has motivated the development of alternative estimators. One popular method is the pre-averaging approach first introduced by Podolskij and Vetter (2009) and further studied by Jacod et al. (2009). The basic underlying idea consists of first averaging out the noise by computing pre-averaged returns and then computing a realized volatility-like estimator using the pre-averaged returns. Although the pre-averaged realized volatility estimator is consistent for integrated volatility, its convergence rate is much slower than that of realized volatility and this can result in finite sample distortions that persist even at very large sample sizes. For this reason, the bootstrap is a useful alternative method of inference in this context.

In this paper, we propose a bootstrap method that can be used to estimate the distribution and the variance of the pre-averaged realized volatility estimator of Jacod et al.

(2009). Our proposal is to resample the pre-averaged returns instead of resampling the original noisy returns. To be valid, the bootstrap needs to mimic the dependence and heterogeneity properties of the (squared) pre-averaged returns. When pre-averaging occurs over overlapping blocks of returns, as in Jacod et al. (2009), the squared pre-averaged returns are k_n-dependent, where k_n denotes the block length of the interval over which the pre-averaging is done and n denotes the sample size. Since k_n is proportional to $\sqrt{n}$, $k_n \to \infty$ as $n \to \infty$, which implies that the pre-averaged returns are strongly dependent. This suggests that a block bootstrap applied to the pre-averaged returns is appropriate and its application amounts to a "blocks of blocks" bootstrap, as proposed by Politis and Romano (1992) and further studied by Bühlmann and Künsch (1995) (see also Künsch (1989)). Nevertheless, as we show here, such a bootstrap scheme is only consistent in our setup when volatility is constant. The reason is that squared pre-averaged returns are heterogenously distributed (in particular, their mean and variance are time-varying) and this creates a bias term in the blocks of blocks bootstrap variance estimator when volatility is stochastic. Thus, to handle both the dependence and heterogeneity of the squared pre-averaged returns, we propose a novel bootstrap approach that combines the wild bootstrap with the blocks of blocks bootstrap. We name this novel approach the wild blocks of blocks bootstrap. Our main contribution is to show that this method consistently estimates the variance and the entire distribution of the pre-averaged estimator of Jacod et al. (2009).

The pre-averaging approach can also be implemented with non-overlapping intervals, as in Podolskij and Vetter (2009). Gonçalves, Hounyo and Meddahi (2013) study the consistency of the wild bootstrap for this estimator. The wild bootstrap exploits the asymptotic independence of the pre-averaged returns when these are computed over non-overlapping intervals. This method is no longer valid when overlapping intervals are used to compute pre-averaged returns since these are strongly dependent. For this reason, a new bootstrap method is needed for the Jacod et al.'s (2009) approach. Although the wild blocks of blocks bootstrap that we propose here requires the choice of an additional tuning parameter (the block size), we suggest an empirical procedure to select the block size that performs well in our simulations.

Other estimators of integrated volatility that are consistent under market microstructure noise include the subsampling approach of Zhang et al. (2005) and the realized kernel estimator of Barndorff-Nielsen et al. (2008) (the maximum likelihood-based estimator of Xiu (2010) is also a recent addition to this literature). The bootstrap could also be useful for inference in the context of these estimators. Indeed, Zhang et al. (2011) showed that the asymptotic normal approximation is often inaccurate for the subsampling realized volatility estimator, whose finite sample distribution is skewed and heavy tailed. They proposed Edgeworth corrections for this estimator as a way to improve upon the standard normal approximation. Similarly, Bandi and Russell (2011) discussed the limitations of asymptotic approximations in the context of realized kernels and proposed an alternative solution. The main reason why we focus on the pre-averaging approach here is that it naturally lends itself to the bootstrap. In particular, we resample the pre-averaged returns instead of the individual returns and exploit the dependence and heterogeneity properties of the pre-averaged returns to prove the consistency of the bootstrap. In addition, the pre-averaging approach has some important advantages compared to the preceding methods, for example it can easily estimate the integrated quarticity or other functionals of volatility.

The rest of this chapter is organized as follows. In the next section, we first introduce the setup, our assumptions and review the existing asymptotic theory of Jacod et al. (2009). Section 2.3 contains the bootstrap results. In Section 2.3.1 we show that the blocks of blocks bootstrap is consistent only when volatility is constant whereas Section 2.3.2 describes the wild blocks of blocks bootstrap and shows its consistency under stochastic volatility and i.i.d. noise. In Section 2.4 we present the Monte Carlo results. Section 2.5 contains an empirical application and Section 2.6 concludes. Two appendices are provided. Appendix C contains the tables with simulation results whereas Appendix D is a mathematical appendix with the proofs.

A word on notation. In this chapter, and as usual in the bootstrap literature, P^* (E^* and Var^*) denotes the probability measure (expected value and variance) induced by the bootstrap resampling, conditional on a realization of the original time series. In addition, for a sequence of bootstrap statistics Z_n^*, we write $Z_n^* = o_{P^*}(1)$ in probability, or $Z_n^* \to^{P^*} 0$, as $n \to \infty$, in probability, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{n\to\infty} P\left[P^*\left(|Z_n^*| > \delta\right) > \varepsilon\right] = 0$. Similarly, we write $Z_n^* = O_{P^*}(1)$ as $n \to \infty$, in probability if for all $\varepsilon > 0$ there exists a $M_\varepsilon < \infty$ such that $\lim_{n\to\infty} P\left[P^*\left(|Z_n^*| > M_\varepsilon\right) > \varepsilon\right] = 0$. Finally, we write $Z_n^* \to^{d^*} Z$ as $n \to \infty$, in probability, if conditional on the sample, Z_n^* weakly converges to Z under P^*, for all samples contained in a set with probability P converging to one.

2.2 Setup, assumptions and review of existing results

2.2.1 Setup and assumptions

Let X denote the latent efficient log-price process defined on a probability space $(\Omega^0, \mathcal{F}^0, P^0)$ equipped with a filtration $(\mathcal{F}_t^0)_{t\geq 0}$. We model X as a Brownian semimartingale process defined by the equation

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s,\ t \geq 0, \tag{2.1}$$

where $a = (a_t)_{t\geq 0}$ is an adapted càdlàg drift process, $\sigma = (\sigma_t)_{t\geq 0}$ is an adapted càdlàg volatility process and $W = (W_t)_{t\geq 0}$ a standard Brownian motion.

The object of interest is the quadratic variation of X, i.e. the process

$$C_t = \int_0^t \sigma_s^2 ds,$$

also known as the integrated volatility. Without loss of generality, we let $t = 1$ and define $C_1 = \int_0^1 \sigma_s^2 ds$ as the integrated volatility of X over a given time interval $[0, 1]$, which we think of as a given day.

The presence of market frictions such as price discreteness, rounding errors, bid-ask spreads, gradual response of prices to block trades, etc, prevent us from observing the true efficient price process X. Instead, we observe a noisy price process Y, observed at time points $t = \frac{i}{n}$ for $i = 0, \ldots, n$, given by

$$Y_t = X_t + \epsilon_t,$$

where ϵ_t represents the noise term that collects all the market microstructure effects.

In order to make both X and Y measurable with respect to the filtration, we define a new probability space $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, P\right)$, which accommodates both processes. To this end, we follow Jacod et al. (2009) and assume one has a second space $\left(\Omega^1, (\mathcal{F}_t^1)_{t\geq 0}, P^1\right)$, where Ω^1 denotes $\mathbb{R}^{[0,1]}$ and $\mathcal{F}^1$ the product Borel-σ-field on Ω^1. Next, let Q_t be a probability measure on $\mathbb{R}$ (Q_t is the marginal distribution of ϵ_t). P^1 denotes the product measure $\otimes_{t\in[0,1]} Q_t$. The filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, P\right)$ on which the process Y lives is then defined with $\Omega = \Omega^0 \times \Omega^1$, $\mathcal{F} = \mathcal{F}^0 \times \mathcal{F}^1$, $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s^0 \times \mathcal{F}_s^1$, and $P = P^0 \otimes P^1$.

We assume that ϵ_t is centered and independent, conditionally on the efficient price process X. In addition, we assume that the conditional variance of ϵ_t is càdlàg. Assumption 1 below collects these assumptions.

Assumption 1.

(i) $E(\epsilon_t | X) = 0$ and ϵ_t and ϵ_s are independent for all $t \neq s$, conditionally on X.

(ii) $\alpha_t = E(\epsilon_t^2 | X)$ is càdlàg and $E(\epsilon_t^8) < \infty$.

Assumption 1 amounts to Assumption (K) in Jacod et al. (2009). As they explain, this assumption is rather general, allowing for time varying variances of the noise and dependence between X and ϵ. See Jacod et al. (2009) for particular examples of market microstructure noise that satisfy Assumption 1.

2.2.2 The pre-averaged estimator and its asymptotic theory

We observe Y at regular time points $\frac{i}{n}$, for $i = 0, \ldots, n$, from which we compute n intraday returns at frequency $\frac{1}{n}$,

$$r_i \equiv Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}, \quad i = 1, \ldots, n.$$

Given that $Y = X + \epsilon$, we can write

$$r_i = \left(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\right) + \left(\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}\right) \equiv r_i^e + \Delta\epsilon_i,$$

where $r_i^e = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ denotes the $\frac{1}{n}$-frequency return on the efficient price process. Under Assumption 1, the order of magnitude of $\Delta\epsilon_i \equiv \epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}$ is $O_P(1)$. In contrast, r_i^e is (conditionally on the path of σ and a) independent and heteroskedastic with (conditional) variance given by $\int_{(i-1)/n}^{i/n} \sigma_s^2 ds$. Thus, its order of magnitude is $O_P\left(n^{-1/2}\right)$. This decomposition shows that the noise completely dominates the observed return process as $n \to \infty$, implying that the usual realized volatility estimator is biased and inconsistent. See Zhang et al. (2005) and Bandi and Russell (2008).

To describe the Jacod et al. (2009) pre-averaging approach, let k_n be a sequence of integers which will denote the window length over which the pre-averaging of returns is done. Similarly, let g be a weighting function on $[0,1]$ such that $g(0) = g(1) = 0$ and $\int_0^1 g(s)^2 ds > 0$, and assume g is continuous and piecewise continuously differentiable with a piecewise Lipschitz derivative g'. An example of a function that satisfies these restrictions is $g(x) = \min(x, 1-x)$.

We introduce the following additional notation. Let

$$\phi_1(s) = \int_s^1 g'(u)\, g'(u-s)\, du \quad \text{and} \quad \phi_2(s) = \int_s^1 g(u)\, g(u-s)\, du,$$

and for $i = 1, 2$, let $\psi_i = \phi_i(0)$. For instance, for $g(x) = \min(x, 1-x)$, we have that $\psi_1 = 1$ and $\psi_2 = 1/12$.

For $i = 0, \ldots, n - k_n + 1$, the pre-averaged returns $\bar{Y}_i$ are obtained by computing the weighted sum of all consecutive $\frac{1}{n}$-horizon returns over each block of size k_n,

$$\bar{Y}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) r_{i+j}.$$

The effect of pre-averaging is to reduce the impact of the noise in the pre-averaged return. Specifically, as shown by Vetter (2008),

$$\bar{X}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(X_{\frac{i+j}{n}} - X_{\frac{i+j-1}{n}}\right) = O_P\left(\sqrt{\frac{k_n}{n}}\right),$$

and

$$\bar{\epsilon}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(\epsilon_{\frac{i+j}{n}} - \epsilon_{\frac{i+j-1}{n}}\right) = O_P\left(\frac{1}{\sqrt{k_n}}\right).$$

Thus, the impact of the noise is reduced the larger k_n is. To get the efficient $n^{-1/4}$ rate of convergence, Jacod et al. (2009) propose to choose a sequence of integers k_n such that the following assumption holds.

Assumption 2. For $\theta \in (0, \infty)$, we have that

$$\frac{k_n}{\sqrt{n}} = \theta + o\left(n^{-1/4}\right). \tag{2.2}$$

This choice implies that the orders of the two terms ($\bar{X}_i$ and $\bar{\epsilon}_i$) are balanced and equal to $O_P\left(n^{-1/4}\right)$. An example that satisfies (2.2) is $k_n = [\theta\sqrt{n}]$.

Based on the pre-averaged returns $\bar{Y}_i$, Jacod et al. (2009) propose the following estimator of integrated volatility,

$$PRV_n = \frac{1}{\psi_2 k_n}\sum_{i=0}^{n-k_n+1} \bar{Y}_i^2 - \frac{\psi_1}{2n\theta^2\psi_2}\sum_{i=1}^{n} r_i^2, \tag{2.3}$$

where ψ_1 and ψ_2 are as defined above.

The first term in (2.3) is an average of realized volatility-like estimators based on pre-averaged returns of length k_n whereas the second term is a bias correction term. As discussed in Jacod et al. (2009), this bias term does not contribute to the asymptotic variance of PRV_n.

In order to give the central limit theorem for PRV_n, we introduce the following numbers that are associated with g,

$$\Phi_{ij} = \int_0^1 \phi_i(s)\,\phi_j(s)\,ds, \quad \text{and} \quad \Psi_{ij} = -\int_0^1 s\phi_i(s)\,\phi_j(s)\,ds.$$

For the simple function $g(x) = \min(x, 1-x)$, $\Phi_{11} = 1/6$, $\Phi_{12} = 1/96$ and $\Phi_{22} = 151/80640$. Under Assumption 1 and (k_n, θ) satisfying (2.2), Jacod et al. (2009) show that as $n \to \infty$,

$$\frac{n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{V}} \to^{st} N(0,1), \tag{2.4}$$

where $\to^{st}$ denotes stable convergence, and

$$V = \frac{4}{\psi_2^2}\int_0^1 \left(\Phi_{22}\theta\sigma_s^4 + 2\Phi_{12}\frac{\sigma_s^2\alpha_s}{\theta} + \Phi_{11}\frac{\alpha_s^2}{\theta^3}\right) ds$$

is the conditional variance of PRV_n. To estimate V consistently, Jacod et al. (2009) propose

$$\begin{aligned}\hat{V}_n &= \frac{4\Phi_{22}}{3\theta\psi_2^4}\sum_{i=0}^{n-k_n+1}\overline{Y}_i^4 + \frac{4}{n\theta^3}\left(\frac{\Phi_{12}}{\psi_2^3} - \frac{\Phi_{22}\psi_1}{\psi_2^4}\right)\sum_{i=0}^{n-2k_n+1}\overline{Y}_i^2\sum_{j=i+k_n}^{i+2k_n-1} r_j^2 \\ &+ \frac{1}{n\theta^3}\left(\frac{\Phi_{11}}{\psi_2^2} - 2\frac{\Phi_{12}\psi_1}{\psi_2^3} + \frac{\Phi_{22}\psi_1^2}{\psi_2^4}\right)\sum_{i=0}^{n-2k_n+1} r_i^2 r_{i+2}^2. \end{aligned} \tag{2.5}$$

Together with the CLT result (2.4), we have that

$$T_n \equiv \frac{n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{\hat{V}_n}} \to^{st} N(0,1).$$

We can use this feasible asymptotic distribution result to build confidence intervals for integrated volatility. In particular, a two-sided feasible $100(1-\alpha)\%$ level interval for $\int_0^1 \sigma_s^2 ds$ is given by:

$$IC_{Feas,1-\alpha} = \left(PRV_n - z_{1-\alpha/2}n^{-1/4}\sqrt{\hat{V}_n}, PRV_n + z_{1-\alpha/2}n^{-1/4}\sqrt{\hat{V}_n}\right),$$

where $z_{1-\alpha/2}$ is such that $\Phi\left(z_{1-\alpha/2}\right) = 1-\alpha/2$, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. For instance, $z_{0.975} = 1.96$ when $\alpha = 0.05$.

2.3 The bootstrap

The goal of this section is to propose a bootstrap method that can be used to consistently estimate the distribution of $n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)$. This justifies the construction of bootstrap percentile confidence intervals for integrated volatility. Although such intervals do not promise asymptotic refinements over confidence intervals based on the asymptotic mixed normal approximation (given by $IC_{Feas,1-\alpha}$), they avoid the need to explicitly estimate the asymptotic variance of the pre-averaged estimator. When the variance estimator is hard to compute (as it is the case here), it is not always clear that estimating the variance is beneficial in small samples. Thus, bootstrap percentile intervals are a very attractive method in these cases.

Gonçalves and Meddahi (2009) proposed bootstrap methods for realized volatility in the absence of market microstructure noise. In their ideal setting, intraday returns r_i are (conditionally on the volatility path) independent, but possibly heteroskedastic due to stochastic volatility, thus motivating the use of a wild bootstrap method.

When intraday returns are contaminated by market microstructure noise, they are no longer conditionally independent, as in Gonçalves and Meddahi (2009). This implies that the wild bootstrap is no longer valid when applied to r_i. Instead, a block bootstrap method applied to the intraday returns would seem appropriate.

One complication arises in this context: the statistic of interest is not symmetric in the observations and the block bootstrap generates blocks of observations that are conditionally independent. In particular, since the first term in PRV_n is an average of the squared pre-averaged returns $\bar{Y}_i^2$, it depends on all the products of intraday returns inside blocks of size k_n. If we generate block bootstrap intraday returns, these will be independent between blocks, implying that the bootstrap statistic may look at many pairs of intraday returns that are independent in the bootstrap world. This not only renders the analysis very complicated but can induce biases in the bootstrap estimator. To avoid this problem when dealing with statistics that are not symmetric in the underlying observations, Künsch (1989), Politis and Romano (1992) and Bühlmann and Künsch (1995) studied the "blocks of blocks" bootstrap, where one applies the block bootstrap to appropriately pre-specified blocks of observations. In our context, the blocks of blocks bootstrap consists of applying a traditional block bootstrap to the squared pre-averaged returns $\bar{Y}_i^2$. As we will see next, this approach is asymptotically valid only when volatility is constant. The reason is that when volatility is stochastic, squared pre-averaged returns are not only dependent but also heterogeneous. The block bootstrap does not capture this heterogeneity unless volatility is constant[1]. In order to capture both the time dependence and the heterogeneity in $\bar{Y}_i^2$, we propose a novel bootstrap procedure that combines the wild bootstrap with the block bootstrap.

Although the consistent estimator of integrated volatility is PRV_n, only the first term in PRV_n drives the variance of the limiting distribution of PRV_n. In particular, as Jacod et al. (2009) have shown, the second term is a bias correction term which does not contribute to the asymptotic variance (it only ensures that the estimator is well centered at the integrated volatility). For this reason, our proposal is to bootstrap only the first

[1]See Gonçalves and White (2002) for a discussion of the impact of mean heterogeneity on the validity of the block bootstrap for the sample mean.

contribution to PRV_n,

$$\widetilde{PRV}_n = \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^2.$$

This statistic depends only on the pre-averaged returns, to which we apply a particular bootstrap scheme. More specifically, let $\left\{\bar{Y}_i^* : i = 0, 1, \ldots, n - k_n + 1\right\}$ denote a bootstrap sample from $\left\{\bar{Y}_i : i = 0, 1, \ldots, n - k_n + 1\right\}$. The bootstrap analogue of $\widetilde{PRV}_n$ is

$$\widetilde{PRV}_n^* = \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^{*2}.$$

Since we do not incorporate a bias correction term in the bootstrap world, we center $\widetilde{PRV}_n^*$ around $E^*\left(\widetilde{PRV}_n^*\right)$. Thus, we use the bootstrap distribution of $n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)$ as an estimator of the distribution of $n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)$.

Next, we consider the blocks of blocks bootstrap approach applied to $\widetilde{PRV}_n$ and show that it is asymptotically invalid when volatility is time-varying. This motivates a new bootstrap method that combines the wild bootstrap with the block bootstrap, which we study in the last subsection.

2.3.1 The blocks of blocks bootstrap

To describe this approach, let $N_n = n - k_n + 2$ denote the total number of pre-averaged returns and let b_n denote the block size. We suppose that $N_n = J_n \cdot b_n$, so that J_n denotes the number of blocks of size b_n one needs to draw to get $N_n = n - k_n + 2$ bootstrap observations. The blocks of blocks bootstrap generates a bootstrap resample $\left\{\bar{Y}_{i-1}^* : i = 1, \ldots, N_n\right\}$ by applying the moving blocks bootstrap of Künsch (1989) to the scaled pre-averaged returns $\left\{\bar{Y}_{i-1} : i = 1, \ldots, N_n\right\}$.

Letting $I_1, \ldots, I_{J_n}$ be i.i.d. random variables distributed uniformly on $\{0, 1, \ldots, N_n - b_n\}$, we set

$$\bar{Y}_{i-1+(j-1)b_n}^* = \bar{Y}_{i-1+I_j} \text{ for } 1 \leq j \leq J_n \text{ and } 1 \leq i \leq b_n.$$

The bootstrap analogue of $\widetilde{PRV}_n$ is

$$\widetilde{PRV}_n^* = \frac{1}{\psi_2 k_n} \sum_{i=1}^{N_n} \bar{Y}_{i-1}^{*2} = \frac{1}{J_n} \sum_{j=1}^{J_n} \left(\frac{1}{b_n} \sum_{i=1}^{b_n} \underbrace{\frac{N_n}{k_n} \frac{1}{\psi_2} \bar{Y}_{I_j+i-1}^2}_{\equiv Z_{I_j+i}} \right),$$

where we let $Z_i \equiv \frac{N_n}{k_n} \frac{1}{\psi_2} \bar{Y}_{i-1}^2$. Note that in our setup, $\bar{Y}_i = \bar{X}_i + \bar{\epsilon}_i = O_P\left(n^{-1/4}\right)$ given that k_n is such that $k_n/\sqrt{n} = \theta + o\left(n^{-1/4}\right)$. This implies that $\bar{Y}_{i-1}^2 = O_P\left(n^{-1/2}\right)$ and therefore $Z_i = \frac{n-k_n+2}{k_n}\frac{1}{\psi_2}\bar{Y}_{i-1}^2$ is $O_P(1)$.

We can easily show that

$$E^*\left(\widetilde{PRV}_n^*\right) = \frac{1}{J_n}\sum_{j=1}^{J_n} E^*\left(\frac{1}{b_n}\sum_{i=1}^{b_n} Z_{I_j+i}\right) = \frac{1}{N_n - b_n + 1}\sum_{j=0}^{N_n-b_n}\left(\frac{1}{b_n}\sum_{i=1}^{b_n} Z_{j+i}\right).$$

Similarly,

$$\begin{aligned} V_n^* &\equiv Var^*\left(n^{1/4}\widetilde{PRV}_n^*\right) = \sqrt{n}E^*\left[\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\frac{1}{b_n}\sum_{i=1}^{b_n}\left(Z_{I_j+i} - E^*\left(\widetilde{PRV}_n^*\right)\right)\right)^2\right] \\ &= \sqrt{n}\frac{1}{J_n}E^*\left(\frac{1}{b_n}\sum_{i=1}^{b_n}\left(Z_{I_1+i} - E^*\left(\widetilde{PRV}_n^*\right)\right)\right)^2 \\ &= \sqrt{n}\frac{b_n}{N_n}\frac{1}{N_n - b_n + 1}\sum_{j=0}^{N_n-b_n}\left(\frac{1}{b_n}\sum_{i=1}^{b_n}\left(Z_{j+i} - E^*\left(\widetilde{PRV}_n^*\right)\right)\right)^2. \end{aligned} \tag{2.6}$$

Our next result studies the convergence of V_n^* when $b_n = (p+1)k_n$, for $p \geq 1$.

Lemma 2.3.1. *Suppose Assumption 1 holds and $k_n \to \infty$ as $n \to \infty$ such that Assumption 2 holds. Let $V_n^* \equiv Var^*\left(n^{1/4}\widetilde{PRV}_n^*\right)$ denote the moving blocks bootstrap variance of $n^{1/4}\widetilde{PRV}_n^*$ based on a block length equal to b_n. Then,*

a) *If $b_n = (p+1)k_n \to \infty$ and $p \geq 1$ is fixed,*

$$p\lim_{n\to\infty} V_n^* = V_p + B_p,$$

where

$$V_p = \int_0^1 \gamma^2(p)_t\, dt$$

with

$$\gamma^2(p)_t = \frac{4}{\psi_2^2}\left[\left(\Phi_{22} + \frac{1}{p+1}\Psi_{22}\right)\theta\sigma_t^4 + 2\left(\Phi_{12} + \frac{1}{p+1}\Psi_{12}\right)\frac{\sigma_t^2\alpha_t}{\theta} + \left(\Phi_{11} + \frac{1}{p+1}\Psi_{11}\right)\frac{\alpha_t^2}{\theta^3}\right],$$

and

$$B_p = \theta(p+1)\left[\int_0^1\left(\sigma_t^2 + \frac{\psi_1}{\theta^2\psi_2}\alpha_t\right)^2 dt - \left(\int_0^1\left(\sigma_t^2 + \frac{\psi_1}{\theta^2\psi_2}\alpha_t\right)dt\right)^2\right].$$

b) When σ is constant, $B_p = 0$ for any $p \geq 1$.

c) If $p \to \infty$ (i.e. $b_n/k_n = p+1 \to \infty$) such that $b_n/n \to 0$, then $V_p \to V \equiv \lim_{n\to\infty} Var\left(n^{1/4}PRV_n\right)$, so that $p\lim_{n\to\infty} V_n^* = V$ if σ is constant and $p\lim_{n\to\infty} V_n^* = \infty$ otherwise.

Part a) of Lemma 2.3.1 shows that when the bootstrap block size b_n is a fixed proportion of the pre-averaging block size k_n, the blocks of blocks bootstrap variance converges in probability to $V_p + B_p$, where B_p is a bias term due to the fact that volatility is time-varying. When σ is constant, B_p is equal to zero for any value of p. If $p \to \infty$ (i.e. if $b_n/k_n \to \infty$ as $n \to \infty$), then $V_p \to V$, the asymptotic variance of $n^{1/4}PRV_n$. Therefore, under this condition and assuming that σ is constant, we obtain the consistency of V_n^* towards V. If σ is stochastic and $p \to \infty$, then V_n^* diverges to infinity since $B_p \to \infty$ as $p \to \infty$.

Lemma 2.3.1 shows that the blocks of blocks bootstrap is consistent for the variance of PRV_n only under constant volatility and if we let the bootstrap block size b_n grow at a faster rate than the pre-averaging block size k_n. This result is related to a consistency result of the blocks of blocks bootstrap established in Bühlmann and Künsch (1995). As they showed, when the statistic of interest is an average of smooth functions of blocks of consecutive stationary strong mixing observations of size k_n, where k_n tends to infinity, the crucial condition for the block bootstrap to be valid is that the block size b_n grows at a faster rate than k_n. This is because the blocks over k_n observations (which in our case correspond to the pre-averaged returns) are strongly dependent for $|i-j| \leq k_n$, where $k_n \to \infty$, and b_n must be large enough to capture this dependence. Bühlmann and Künsch (1995) consider observations generated from a stationary strong mixing process and therefore they do not find any bias problem related to heterogeneity. Nevertheless, this becomes a problem in our context when volatility is stochastic. Therefore, a different bootstrap method is required to handle both the time dependence and the heterogeneity of pre-averaged returns.

2.3.2 The wild blocks of blocks bootstrap

In this section, we propose and study the consistency of a novel bootstrap method for pre-averaged returns based on overlapping blocks of k_n intraday returns. It combines the blocks of blocks bootstrap with the wild bootstrap and in this manner gets rid of the bias term B_p associated with the blocks of blocks bootstrap variance V_n^* in (2.6).

As in the previous section, for $p \geq 1$, let $b_n = (p+1)\,k_n$, and assume that J_n is such that $J_n \cdot b_n = N_n$. Let $\eta_1, \ldots, \eta_{J_n}$ be i.i.d. random variables whose distribution is independent of the original sample. Denote by $\mu_q^* = E^*\left(\eta_j^q\right)$ its q-th order moments. For

$j = 1, \ldots, J_n$, let

$$\bar{B}_j = \frac{1}{b_n} \sum_{i=1}^{b_n} \bar{Y}^2_{i-1+(j-1)b_n}$$

denote the block average of the squared pre-averaged returns $\bar{Y}^2_{i-1+(j-1)b_n}$ for block j. We then generate the bootstrap pre-averaged squared returns as follows,

$$\bar{Y}^{*2}_{i-1+(j-1)b_n} = \bar{B}_{j+1} + \left(\bar{Y}^2_{i-1+(j-1)b_n} - \bar{B}_{j+1} \right) \eta_j, \text{ for } 1 \leq j \leq J_n - 1 \text{ and for } 1 \leq i \leq b_n. \tag{2.7}$$

For the last block $j = J_n$, $\bar{B}_{j+1}$ is not available and therefore we let

$$\bar{Y}^{*2}_{i-1+(j-1)b_n} = \bar{B}_j + \left(\bar{Y}^2_{i-1+(j-1)b_n} - \bar{B}_j \right) \eta_j, \text{ for } 1 \leq i \leq b_n. \tag{2.8}$$

Our method is related to the wild bootstrap approach of Wu (1986) and Liu (1988). More specifically, in Wu (1986) and Liu (1988), the statistic of interest is $\bar{X}_n$, where X_i is independently but heterogeneously distributed with mean μ_i and variance σ_i^2. Their wild bootstrap generates X_i^* as

$$X_i^* = \bar{X}_n + \left(X_i - \bar{X}_n \right) \eta_i, \text{ for } 1 \leq i \leq n,$$

where η_i is i.i.d. $(0, 1)$. Liu (1988) shows that the bootstrap distribution of $\sqrt{n} \left(\bar{X}_n^* - \bar{X}_n \right)$ is consistent for the distribution of $\sqrt{n} \left(\bar{X}_n - \bar{\mu}_n \right)$, where $\bar{\mu}_n = n^{-1} \sum_{i=1}^n \mu_i$, provided $\frac{1}{n} \sum_{i=1}^n \left(\mu_i - \bar{\mu}_n \right)^2 \to 0$ (and some other regularity conditions).

Our bootstrap method can be seen as a generalization of the wild bootstrap of Wu (1986) and Liu (1988) to the k_n-dependent case. In particular, here the statistic of interest is an average of blocks of observations of size k_n,

$$\widetilde{PRV}_n = \frac{1}{N_n} \sum_{i=1}^{N_n} Z_i,$$

where $Z_i \equiv \frac{N_n}{k_n} \frac{1}{\psi_2} \bar{Y}^2_{i-1}$ has time-varying moments and is k_n-dependent (conditionally on X), i.e. Z_i is independent of Z_j for all $|i - j| > k_n$.

To preserve the serial dependence, we divide the data into J_n non-overlapping blocks of size b_n and generate the bootstrap observations within a given block j using the same external random variable η_j. This preserves the dependence within each block. When there is no dependence, we can take $b_n = 1$, in which case our bootstrap method amounts to Liu's wild bootstrap with one difference: instead of centering each bootstrap observation Z_i^* around the overall mean $\widetilde{PRV}_n$, we center Z_i^* around Z_{i+1}. The reason for the new centering is that μ_i in our context does not satisfy Liu's condition $\frac{1}{n} \sum_{i=1}^n \left(\mu_i - \bar{\mu}_n \right)^2 \to 0$

(unless volatility is constant). Hence centering around $\widetilde{PRV}_n$ does not work here. Instead, we show that centering around Z_{i+1} yields an asymptotically valid bootstrap method for $\widetilde{PRV}_n$ even when volatility is stochastic.

The bootstrap data generating process (2.7) and (2.8) yields a bootstrap sample $\left\{\bar{Y}_0^{*2}, \ldots, \bar{Y}_{N_n-1}^{*2}\right\}$ which we use to compute

$$PRV_n^* = \frac{1}{\psi_2 k_n} \sum_{i=1}^{N_n} \bar{Y}_{i-1}^{*2},$$

the wild blocks of blocks bootstrap analogue of $\widetilde{PRV}_n$. Let

$$\bar{B}_j^* = \frac{1}{b_n} \sum_{i=1}^{b_n} \bar{Y}_{i-1+(j-1)b_n}^{*2}$$

be the bootstrap analogue of $\bar{B}_j$. Given (2.7), we have that for $j = 1, \ldots, J_n - 1$,

$$\bar{B}_j^* = \bar{B}_{j+1} + \left(\bar{B}_j - \bar{B}_{j+1}\right) \eta_j,$$

whereas from (2.8),, $\bar{B}_j^* = \bar{B}_j$ for $j = J_n$. This implies that we can write

$$\begin{aligned} PRV_n^* &= \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n} \frac{1}{b_n} \sum_{i=1}^{b_n} \bar{Y}_{i-1+(j-1)b_n}^{*2} = \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n-1} \bar{B}_j^* + \frac{b_n}{\psi_2 k_n} \bar{B}_{J_n}^* \\ &= \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n-1} \left[\bar{B}_{j+1} + \left(\bar{B}_j - \bar{B}_{j+1}\right) \eta_j\right] + \frac{b_n}{\psi_2 k_n} \bar{B}_{J_n}. \end{aligned}$$

We can now easily obtain the bootstrap mean and variance of PRV_n^*. In particular,

$$E^*\left(PRV_n^*\right) = \frac{b_n}{\psi_2 k_n} \left(\sum_{j=1}^{J_n-1} \bar{B}_{j+1} + \bar{B}_{J_n}\right) + \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n-1} \left(\bar{B}_j - \bar{B}_{j+1}\right) E^*\left(\eta_j\right),$$

and

$$V_n^* \equiv Var^*\left(n^{1/4} PRV_n^*\right) = \frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \left(\bar{B}_j - \bar{B}_{j+1}\right)^2 Var^*\left(\eta_j\right).$$

Our next result studies the convergence of V_n^* when $b_n = (p+1) k_n$ and p is either fixed such that $p \geq 1$, or $p \to \infty$.

Lemma 2.3.2. *Suppose Assumption 1 holds and $k_n \to \infty$ as $n \to \infty$ such that Assumption 2 holds. Let $V_n^* \equiv Var^*\left(n^{1/4} PRV_n^*\right)$ denote the wild blocks of blocks bootstrap variance of $n^{1/4} PRV_n^*$ based on a block length equal to b_n and external random variables $\eta_j \sim$ i.i.d. with mean $E^*\left(\eta_j\right)$ and variance $Var^*\left(\eta_j\right)$. Then,*

a) *If* $b_n = (p+1)\, k_n \to \infty$ *and* p *is fixed,*

$$p \lim_{n\to\infty} V_n^* = 2Var^*(\eta_j)\, V_p + O_P\left(\frac{1}{p}\right),$$

where V_p *is as defined in Lemma 2.3.1.*

b) If $p \to \infty$ (i.e. $b_n/k_n = p+1 \to \infty$) such that $b_n/n \to 0$ and $Var^*(\eta_j) = 1/2$, then $V_p \to V \equiv \lim_{n\to\infty} Var\left(n^{1/4} PRV_n\right)$ so that $p\lim_{n\to\infty} V_n^* = V$.

This result shows that if we let b_n grow faster than k_n and $Var^*(\eta_j) = 1/2$, the wild blocks bootstrap variance estimator is consistent for the asymptotic variance of PRV_n under Assumptions 1 and 2. Given the consistency of the bootstrap variance estimator, we can now prove the consistency of the bootstrap distribution of $n^{1/4}(PRV_n^* - E^*(PRV_n^*))$.

Theorem 2.3.1. *Suppose Assumption 1 holds and* $k_n \to \infty$ *as* $n \to \infty$ *such that Assumption 2 holds. Let* PRV_n^* *be the pre-averaged realized volatility estimator based on a block length equal to* b_n *and an external random variable* $\eta_j \sim$ *i.i.d.* $(E^*(\eta_j), Var^*(\eta_j))$ *such that* $Var^*(\eta_j) = \frac{1}{2}$*, and for any* $\delta > 0$ $E^*|\eta_j|^{2+\delta} \le \Delta < \infty$*. If* b_n *is such that* $b_n = (p+1)\, k_n$*,* $b_n/n \to 0$ *and* $p \to \infty$*, then*

$$\sup_{x\in\mathbb{R}} \left| P^*\left(n^{1/4}(PRV_n^* - E^*(PRV_n^*)) \le x\right) - P\left(n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right) \le x\right)\right| \to^P 0 \text{ as } n \to \infty.$$

2.4 Monte Carlo results

In this section, we compare the finite sample performance of the bootstrap with the feasible asymptotic theory for confidence intervals of integrated volatility.

We consider two data generating processes in our simulations. First, following Zhang et al. (2005), we use the one-factor stochastic volatility (SV1F) model of Heston (1993) as our data-generating process, i.e.

$$dX_t = (\mu - \nu_t/2)\, dt + \sigma_t dB_t,$$

and

$$d\nu_t = \kappa(\alpha - \nu_t)\, dt + \gamma(\nu_t)^{1/2}\, dW_t,$$

where $\nu_t = \sigma_t^2$, and we assume $Corr(B, W) = \rho$. The parameter values are all annualized. In particular, we let $\mu = 0.05/252$, $\kappa = 5/252$, $\alpha = 0.04/252$, $\gamma = 0.05/252$, $\rho = -0.5$. The size of the market microstucture noise is an important parameter. We follow Barndorff-Nielsen et al. (2009) and model the noise magnitude as $\xi^2 = \omega^2/\sqrt{\int_0^1 \sigma_s^4 ds}$. We fix ξ^2

equal to 0.0001, 0.001 and 0.01 and let $\omega^2 = \xi^2 \sqrt{\int_0^1 \sigma_s^4 ds}$. These values are motivated by the empirical study of Hansen and Lunde (2006), who investigate 30 stocks of the Dow Jones Industrial Average.

We also consider the two-factor stochastic volatility (SV2F) model analyzed by Barndorff-Nielsen et al. (2009), where [2]

$$\begin{aligned}
dX_t &= \mu dt + \sigma_t dB_t, \\
\sigma_t &= s\text{-}exp\left(\beta_0 + \beta_1 \tau_{1t} + \beta_2 \tau_{2t}\right), \\
d\tau_{1t} &= \alpha_1 \tau_{1t} dt + dB_{1t}, \\
d\tau_{2t} &= \alpha_2 \tau_{2t} dt + \left(1 + \phi \tau_{2t}\right) dB_{2t}, \\
corr\left(dW_t, dB_{1t}\right) &= \varphi_1, \ corr\left(dW_t, dB_{2t}\right) = \varphi_2.
\end{aligned}$$

We follow Huang and Tauchen (2005) and set $\mu = 0.03$, $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.00137$, $\alpha_2 = -1.386$, $\phi = 0.25$, $\varphi_1 = \varphi_2 = -0.3$. We initialize the two factors at the start of each interval by drawing the persistent factor from its unconditional distribution, $\tau_{10} \sim N\left(0, \frac{-1}{2\alpha_1}\right)$, and by starting the stronlgly mean-reverting factor at zero.

We simulate data for the unit interval $[0, 1]$ and normalize one second to be 1/23400, so that $[0, 1]$ is thought to span 6.5 hours. The observed Y process is generated using an Euler scheme. We then construct the $\frac{1}{n}$-horizon returns $r_i \equiv Y_{i/n} - Y_{(i-1)/n}$ based on samples of size n.

We use two different values of θ: $\theta = 1/3$, as in Jacod et al. (2009), and $\theta = 1$, as in Christensen, Kinnebrock and Podolskij (2010). The latter value corresponds to a conservative choice of k_n. We also follow the literature and use the weight function $g(x) = \min(x, 1 - x)$ to compute the pre-averaged returns.

In order to reduce finite sample biases associated with Riemann integrals, we follow Jacod et al. (2009) and Hautsch and Podolskij (2012) and use the finite sample adjustments version of the pre-averaged realized volatility estimator,

$$PRV_n^a = \left(1 - \frac{\psi_1^{k_n}}{2n\theta^2 \psi_2^{k_n}}\right)^{-1} \left(\frac{n}{n - k_n + 2} \frac{1}{\psi_2^{k_n} k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^2 - \frac{\psi_1^{k_n}}{2n\theta^2 \psi_2^{k_n}} \sum_{i=1}^{n} r_i^2\right),$$

where $\psi_1^{k_n} = k_n \sum_{i=1}^{k_n} \left(g\left(\frac{i}{k_n}\right) - g\left(\frac{i-1}{k_n}\right)\right)^2$ and $\psi_2^{k_n} = \frac{1}{k_n} \sum_{i=1}^{k_n} g^2\left(\frac{i}{k_n}\right)$. Similarly, $\hat{V}_n$ as defined

[2] The function s-exp is the usual exponential function with a linear growth function splined in at high values of its argument: $s\text{-}exp(x) = exp(x)$ if $x \leq x_0$ and $s\text{-}exp(x) = \frac{exp(x_0)}{\sqrt{x_0 - x_0^2 + x^2}}$ if $x > x_o$, with $x_0 = log(1.5)$.

in (2.5) replaces Φ_{11}, Φ_{12} and Φ_{22} by their Riemann approximations,

$$\Phi_{11}^{k_n} = k_n \left(\sum_{i=1}^{k_n} \left(\phi_1^{k_n}(j) \right)^2 - \frac{1}{2} \left(\phi_1^{k_n}(0) \right)^2 \right), \quad \Phi_{12}^{k_n} = \frac{1}{k_n} \left(\sum_{i=1}^{k_n} \phi_1^{k_n}(j) \, \phi_2^{k_n}(j) - \frac{1}{2} \phi_1^{k_n}(0) \, \phi_2^{k_n}(0) \right), \quad \text{and}$$

$$\Phi_{22}^{k_n} = \frac{1}{k_n^3} \left(\sum_{i=1}^{k_n} \left(\phi_2^{k_n}(j) \right)^2 - \frac{1}{2} \left(\phi_2^{k_n}(0) \right)^2 \right),$$

where

$$\phi_1^{k_n}(j) = k_n \sum_{i=j+1}^{k_n-1} \left(g\left(\frac{i-1}{k_n} \right) - g\left(\frac{i}{k_n} \right) \right) \left(g\left(\frac{i-j-1}{k_n} \right) - g\left(\frac{i-j}{k_n} \right) \right), \quad \text{and}$$

$$\phi_2^{k_n}(j) = \sum_{i=j+1}^{k_n-1} g\left(\frac{i}{k_n} \right) - g\left(\frac{i-j}{k_n} \right).$$

Tables 3.4 and 3.5 give the actual rates of 95% confidence intervals of integrated volatility for the SV1F and the SV2F models, respectively, computed over 10,000 replications. Results are presented for eight different samples sizes: $n = 23400$, 11700, 7800, 4680, 1560, 780, 390 and 195, corresponding to "1-second", "2-second", "3-second", "5-second", "15-second", "30-second", "1-minute" and "2-minute" frequencies.

In our simulations, bootstrap intervals use 999 bootstrap replications for each of the 10,000 Monte Carlo replications. We consider the bootstrap percentile method computed at the 95% level. To generate the bootstrap data we use the following external random variables $\eta_j \sim$ i.i.d. $N(0, 1/2)$. The choice of the bootstrap block size is critical. We follow Politis, Romano and Wolf (1999) and use the Minimum Volatility Method to choose the bootstrap block. Details of the algorithm are given in Appendix C.

For the two models, all intervals tend to undercover. The degree of undercoverage is especially large for smaller values of n, when sampling is not too frequent. The SV2F model exhibits overall larger coverage distortions than the SV1F model, for all sample sizes. Results are sensitive to the value of the tuning parameter θ. When $\theta = 1/3$, larger market microstructure effects induce larger coverage distortions. In particular, the coverage distortions are very important when $\xi^2 = 0.01$ in comparison to the case where market microstructure effects are moderate or negligible ($\xi^2 = 0.001$ and $\xi^2 = 0.0001$). This reflects the fact that for this value of θ, k_n is not sufficiently large to allow pre-averaging to remove the market microstructure bias. The pre-averaged estimator is biased in finite samples and this explains the finite sample distortions. In contrast, for the conservative choice of k_n, results are not very sensitive to the noise magnitude. The reason is that the larger is the block size over which the pre-averaging is done, the smaller is the impact of the noise.

In all cases, the bootstrap outperforms the existing first order asymptotic theory. As expected, the average chosen block size is larger for larger sample sizes, but our results show that it is not sensitive to the noise magnitude. This is because the noise magnitude is almost irrelevant for the intensity of the serial autocorrelation of the square pre-averaged returns (as confirmed by simulations not reported here).

2.5 Empirical results

In this section, we implement the wild blocks of blocks bootstrap on high frequency data and compare it to the existing feasible asymptotic procedure of Jacod et al. (2009). The data consists of transaction log prices of General Electric (GE) shares carried out on the New York Stock Exchange (NYSE) in October 2011. Our procedure for cleaning the data is exactly identical to that used by Barndorff-Nielsen et al. (2008) (for further details see Barndorff-Nielsen et al. (2009)). For each day, we consider data from the regular exchange opening hours from time stamped between 9:30 a.m. until 4 p.m.

We implement the pre-averaged realized volatility estimator of Jacod et al. (2009) on returns recorded every S transactions, where S is selected each day so that there are approximately 1493 observations a day. This means that on average these returns are recorded roughly every 15 seconds. Table 3.6 in Appendix C provides the number of transactions per day and the sample size for the pre-averaged returns.

To implement the pre-averaged realized volatility estimator, we select the tunning parameter θ by following the conservative rule ($\theta = 1$, implying that $k_n = \sqrt{n}$). To choose the block size b_n, we follow Politis, Romano and Wolf (1999) and use the Minimum Volatility Method (see Appendix C for details).

Figure 3.2 in Appendix C shows daily 95% confidence intervals (CIs) for integrated volatility using both methods, the wild blocks of blocks bootstrap and the existing feasible asymptotic procedure of Jacod et al. (2009). The confidence intervals based on the bootstrap method are usually wider than the confidence intervals using the feasible asymptotic theory.[3] This is especially true in periods with large volatility. To gain further insight on the behavior of our intervals for these periods, we implemented the test for jumps of Barndorff- Nielsen and Shephard (2006) using a moderate sample size (2-minute sampling intervals). It turns out that these days often correspond to days on which there is evidence for jumps (in particular for the 13, 17, 20 and 26 of October 2011). Since neither

[3] Nevertheless, as our Monte Carlo simulations showed, the latter typically have undercoverage problems whereas the bootstrap intervals have coverage rates closer to the desired level. Therefore if the goal is to control the coverage probability, shorter intervals are not necessarily better. The figures also show a lot of variability in the daily estimate of integrated volatility.

of the two types of intervals are valid in the presence of jumps, further analysis should be pursued for these particular days. In particular, we should rely on estimation methods that are robust to jumps such as the pre-averaged multipower variation method proposed by Podolskij and Vetter (2009) or the quantile estimation method of Christensen, Oomen, and Podolskij (2010).

2.6 Conclusion

In this chapter we propose the bootstrap as a method of inference for integrated volatility in the context of the pre-averaged realized volatility estimator proposed by Jacod et al. (2009). We show that the “blocks of blocks" bootstrap method suggested by Politis and Romano (1992) is valid in this context only when volatility is constant. This is due to the heterogeneity of the squared pre-averaged returns when volatility is stochastic.

To simultaneously handle the dependence and heterogeneity of the pre-averaged returns, we propose a novel bootstrap procedure that combines the wild and the blocks of blocks bootstrap. We provide a set of conditions under which this method is asymptotically valid to first order. Our Monte Carlo simulations show that the wild blocks of blocks bootstrap improves the finite sample properties of the existing first order asymptotic theory. The empirical results suggest that this bootstrap method is generally more accurate than the existing feasible approach of Jacod et al. (2009). In future work, we plan to generalize the wild blocks of blocks bootstrap for inference on multivariate integrated volatility as considered by Christensen, Kinnebrock and Podolskij (2010). Bootstrap variance-covariances matrices are naturally positive semi-definite, which is very important for empirical applications.

Chapter 3

Bootstrapping realized covolatility measures under local Gaussianity assumption

3.1 Introduction

Realized measures of volatility have become extremely popular in the last decade as higher and higher frequency returns are available. Despite the fact that these statistics are measured over large samples, their finite sample distributions are not necessarily well approximated by their asymptotic mixed-Gaussian distributions. This is especially true for realized statistics that are not robust to market microstructure noise since in this case researchers usually face a trade-off between using large sample sizes and incurring in market microstructure biases. This has spurred interest in developing alternative approximations based on the bootstrap. In particular, Gonçalves and Meddahi (2009) have recently proposed bootstrap methods for realized volatility whereas Dovonon, Gonçalves and Meddahi (2013) have studied the application of the bootstrap in the context of realized regressions.

The main contribution of this chapter is to propose a new bootstrap method that exploits the local Gaussianity framework described in Mykland and Zhang (2009, 2011). As these authors explain, one useful way of thinking about inference in the context of realized measures is to assume that returns have constant variance and are conditionally Gaussian over blocks of consecutive M observations. Roughly speaking, a high frequency return of a given asset is equal in law to the product of its volatility (the spot volatility) multiplied by a normal standard distribution. Mykland and Zhang (2009) show that this local Gaussianity assumption is useful in deriving the asymptotic theory for the estimators used in this literature by providing an analytic tool to find the asymptotic behaviour without calculations being too cumbersome. This approach also has the advantage of yielding more efficient estimators by varying the size of the block (see Mykland and Zhang (2009) and Mykland, Shephard and Sheppard (2012).

The main idea of this chapter is to see how and to what extent this local Gaussianity assumption can be explored to generate a bootstrap approximation. In particular, we propose and analyze a new bootstrap method that relies on the conditional local Gaussianity of intraday returns. The new method (which we term the local Gaussian bootstrap) consists of dividing the original data into non-overlapping blocks of M observations and then generating the bootstrap observations at each frequency within a block by drawing a random draw from a normal distribution with mean zero and variance given by the realized volatility over the corresponding block. Using Mykland and Zhang's (2009) blocking approach, one can act as if the instantaneous volatility is constant over a given block of consecutive observations. In practice, the volatility of asset returns is highly persistent, especially over a daily horizon, implying that it is at least locally nearly constant.

We focus on two realized measures in this chapter: realized volatility and realized regression coefficients. The latter can be viewed as a smooth function of the realized covariance matrix. Our proposal in this case is to generate bootstrap observations on the vector that collects the intraday returns that enter the regression model by applying the same idea as in the univariate case. Specifically, we generate bootstrap observations on the vector of variables of interest by drawing a random vector from a multivariate normal distribution with mean zero and covariance matrix given by the realized covariance matrix computed over the corresponding block.

Our findings for realized volatility are as follows. When M is fixed, the local Gaussian bootstrap is asymptotically correct but it does not offer any asymptotic refinements. More specifically, the first four bootstrap cumulants of the t-statistic based on realized volatility and studentized with a variance estimator that is based on a block size of M do not match the cumulants of the original t-statistic to higher order (although they are consistent). Note that when $M = 1$, the new bootstrap method coincides with the wild bootstrap of Gonçalves and Meddahi (2009) based on a $N(0,1)$ external random variable. As Gonçalves and Meddahi (2009) show, this is not an optimal choice, which is in line with our results. Therefore, our result generalizes that of Gonçalves and Meddahi (2009) to the case of a fixed $M > 1$. However, if the block length $M \to \infty$ at rate $o(h^{-1/2})$ (where h^{-1} denotes the sample size), then the local Gaussian bootstrap is able to provide an asymptotic refinement. In particular, we show that the first and third bootstrap cumulants of the t-statistic converge to the corresponding cumulants at the rate $o\left(h^{-1/2}\right)$, which implies that the local Gaussian bootstrap offers a second-order refinement. In this case, the local Gaussian bootstrap is an alternative to the optimal two-point distribution wild bootstrap proposed by Gonçalves and Meddahi (2009). More interestingly, we also show that the local Gaussian bootstrap is able to match the second and fourth order cumulants through order $o(h)$, which implies that this method is able to provide a third-order asymptotic refinement. This is contrast to the optimal wild bootstrap methods of Gonçalves and Meddahi (2009), which can not deliver third-order asymptotic refinements.

For the realized regression estimator proposed by Mykland and Zhang (2009), the local Gaussian bootstrap matches the cumulants of the t-statistics through order $o\left(h^{-1/2}\right)$ when $M \to \infty$ at rate $o\left(h^{-1/2}\right)$. Thus, this method can promise second-order refinements. This is contrast with the pairs bootstrap studied by Dovonon, Gonçalves and Meddahi (2013), which is only first-order correct.

Our Monte Carlo simulations suggest that the new bootstrap method we propose improves upon the first-order asymptotic theory in finite samples and outperforms the existing bootstrap methods.

The rest of this chapter is organized as follows. In the next section, we first introduce the setup, our assumptions and describe the local Gaussian bootstrap. In Sections 3.3 and 3.4 we establish the consistency of this method for realized volatility and realized betas, respectively. Section 3.5 contains the higher-order asymptotic properties of the bootstrap cumulants. Section 3.6 contains simulations, Section 3.7 contains one empirical application and Section 3.8 concludes. Three appendices are provided. Appendix E contains the tables with simulation results whereas Appendix F and Appendix G contain the proofs.

3.2 Framework and the local Gaussian bootstrap

The statistics of interest in this chapter can be written as smooth functions of the realized multivariate volatility matrix. Here we describe the theoretical framework for multivariate high frequency returns and introduce the new bootstrap method we propose. Sections 3.3 and 3.4 will consider in detail the theoretical properties of this method for the special cases of realized volatility and realized beta, respectively.

We follow Mykland and Zhang (2009) and assume that the log-price process $X_t = \left(X_t^{(1)} \cdots X_t^{(d)}\right)'$ of a d-dimensional vector of assets is defined on a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$. We model X as a Brownian semimartingale process that follows the equation,

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s,\ t \geq 0, \tag{3.1}$$

where $\mu = (\mu_t)_{t\geq 0}$ is a d-dimensional predictable locally bounded drift vector, $\sigma = (\sigma_t)_{t\geq 0}$ is an adapted càdlàg $d \times d$ locally bounded spot covolatility matrix and $W = (W_t)_{t\geq 0}$ is d-dimensional Brownian motion.

We follow Barndorff-Nielsen et al. (2006) and assume that the spot covariance matrix $\Sigma_t = \sigma_t \sigma_t'$ is invertible and satisfies the following assumption

$$\Sigma_t = \Sigma_0 + \int_0^t a_s ds + \int_0^t b_s dW_s + \int_0^t v_s dZ_s, \tag{3.2}$$

where a, b, and v are all adapted càdlàg processes, with a also being predictable and locally bounded, and Z is a vector Brownian motion independent of W.

The representation in (3.1) and (3.2) is rather general as it allows for leverage and drift effects. Assumption 2 of Mykland and Zhang (2009) or equation (1) of Mykland and Zhang (2011) also impose a Brownian semimartingale structure on the instantaneous covariance matrix Σ. Equation (3.2) rules out jumps in volatility, but this can be relaxed (see Assumption H1 of Barndorff-Nielsen et al. (2006) for a weaker assumption on Σ).

Suppose we observe X over a fixed time interval $[0, 1]$ at regular time points ih, for $i = 0, \ldots, 1/h$, from which we compute $1/h$ intraday returns at frequency h,

$$y_i \equiv X_{ih} - X_{(i-1)h} = \int_{(i-1)h}^{ih} \mu_t dt + \int_{(i-1)h}^{ih} \sigma_t dW_t, \quad i = 1, \ldots, \frac{1}{h}, \tag{3.3}$$

where we will let y_{ki} to denote the i-th intraday return on asset k , $k = 1, \ldots, d$.

As equation (3.3) shows, the intraday returns y_i depend on the drift μ, unfortunately when carrying out inference for observations in a fixed time interval the process μ_t cannot be consistently estimated. For most purposes it is only a nuisance parameter. To deal with this, Mykland and Zhang (2009) propose to work with a new probability measure which is measure theoretically equivalent to P and under which there is no drift (a statistical risk neutral measure). They pursue the analysis further and propose an approximation measure Q_h defined on the discretized observations X_{ih} only, for which the volatility is constant on each of the $\frac{1}{Mh}$ non overlapping blocks of size M. Since M is the number of high frequency returns within a block, we have that $M \leq \frac{1}{h}$.

Specifically, under the approximate measure Q_h, in each block $j = 1, \ldots, \frac{1}{Mh}$, we have,

$$y_i = \frac{1}{\sqrt{M}} C_{(j)} \eta_{i+(j-1)M}, \ \forall i \in \left((j-1) M, jM \right], \tag{3.4}$$

where $\eta_{i+(j-1)M} \sim i.i.d. N(0, I_d)$, I_d is a $d \times d$ identity matrix and $C_{(j)} = \sqrt{Mh}\sigma_{(j-1)Mh}$, where $C_{(j)}$ is such that $C_{(j)} C_{(j)}' = \Gamma_{(j)} \equiv \int_{(j-1)Mh}^{jMh} \Sigma_u du$ (see Mykland and Zhang (2009), p.1417 for a formal definition of Q_h).

The true distribution is P, but we prefer to work with Q_h since then calculations are much simpler. Afterwards we adjust results back to P using the likelihood ratio (Radon-Nikodym derivative) dQ_h/dP.

Remark 1. As pointed out in Mykland and Zhang's (2009) Theorem 3 and, in Mykland and Zhang's (2011) Theorem 1, the measure P and its approximation Q_h are contiguous on the observables. This is to say that for any sequence $\mathcal{A}_h$ of sets, $P(\mathcal{A}_h) \to 0$ if and only if $Q_h(\mathcal{A}_h) \to 0$ (see Mykland and Zhang (2012) p. 169 for more details). In particular, if an estimator is consistent under Q_h, it is also consistent under P. Rates of convergence (typically $h^{-1/2}$) are also preserved, but the asymptotic distribution may change (for instances of this, see Examples 3 and 5 of Mykland and Zhang (2009). More specifically, when adjusting from Q_h to P, the asymptotic variance of the estimator is unchanged (due to the preservation of quadratic variation under limit operations), while the asymptotic bias may change (see Remark 4 of Mykland and Zhang (2009)). It appears that a given sequence Z_h of martingales will have exactly the same asymptotic distribution under Q_h and P, when the Q_h martingale part of the log likelihood ratio $log(dP/dQ_h)$ has zero asymptotic covariation with Z_h. In this case, we do not need to adjust the distributional result from Q_h to P. Two important examples where this is true are the realized volatility and realized beta which we will study in details in Sections 3 and 4.

Remark 2. In the particular case where the window length M increases with the sample size h^{-1} at rate $o(h^{-1/2})$, there is also no contiguity adjustment (see Remark 2 of Mykland and Zhang (2011)).

Next we introduce a new bootstrap method that exploits the structure of (3.4). In particular, we mimic the original observed vector of returns, and we use the normality of the data and replace $C_{(j)}$ by its estimate $\hat{C}_{(j)}$, where $\hat{C}_{(j)}$ is such that $\hat{C}_{(j)}\hat{C}'_{(j)} = \sum_{i=1}^{M} y_{i+(j-1)M} y'_{i+(j-1)M} = \hat{\Gamma}_{(j)}$. That is, we follow the main idea of Mykland and Zhang (2009), and assume constant volatility within blocks. Then, inside each block j of size M ($j = 1, \ldots, \frac{1}{Mh}$), we generate the M vector of returns as follows,

$$y^*_{i+(j-1)M} = \frac{1}{\sqrt{M}} \hat{C}_{(j)} \eta_{i+(j-1)M},\ 1 = 1, \ldots, M, \tag{3.5}$$

where $\eta_{i+(j-1)M} \sim i.i.d.N(0, I_d)$ across (i, j), and I_d is a $d \times d$ identity matrix.

In this chapter, and as usual in the bootstrap literature, P^* (E^* and Var^*) denotes the probability measure (expected value and variance) induced by the bootstrap resampling, conditional on a realization of the original time series. In addition, for a sequence of bootstrap statistics Z_h^*, we write $Z_h^* = o_{P^*}(1)$ in probability, or $Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under P, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{h\to 0} P\left[P^*\left(|Z_h^*| > \delta\right) > \varepsilon\right] = 0$. Similarly, we write $Z_h^* = O_{P^*}(1)$ as $h \to 0$, in probability if for all $\varepsilon > 0$ there exists a $M_\varepsilon < \infty$ such that $\lim_{h\to 0} P\left[P^*\left(|Z_h^*| > M_\varepsilon\right) > \varepsilon\right] = 0$. Finally, we write $Z_h^* \to^{d_*} Z$ as $h \to 0$, in probability under P, if conditional on the sample, Z_h^* weakly converges to Z under P^*, for all samples contained in a set with probability converging to one.

The following result is crucial in obtaining our bootstrap results.

Theorem 3.2.1. *Let Z_h^* be a sequence of bootstrap statistics. Given the probability measure P and its approximation Q_h, we have that*

$Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under P, if and only if $Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under Q_h.

Proof of Theorem 3.2.1 For any $\varepsilon > 0$, $\delta > 0$, letting $\mathcal{A}_h \equiv \left\{P^*\left(|Z_h^*| > \delta\right) > \varepsilon\right\}$, we have that

$Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under P, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{h\to 0} P\left(\mathcal{A}_h\right) = 0$. This is equivalent to $\lim_{h\to 0} Q_h\left(\mathcal{A}_h\right) = 0$, since P and Q_h are contiguous (see Remark 1). It follows then that $Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under Q_h. The inverse follows similarly.

Theorem 3.2.1 provides a theoretical justification to derive bootstrap consistency results under the approximation measure Q_h as well as under P. This simplifies the bootstrap inference. We will subsequently rely on this theorem to establish the bootstrap consistency results.

3.3 Results for realized volatility

3.3.1 Existing asymptotic theory

To describe the asymptotic properties of realized volatility, we need to introduce some notation. For any $q > 0$, define the realized q-th order power variation (cf. Remark 8 of Mykland and Zhang (2009)) as

$$R_q \equiv Mh \sum_{j=1}^{1/Mh} \left(\frac{RV_{j,M}}{Mh}\right)^{q/2}.$$

where $RV_{j,M} = \sum_{i=1}^{M} y_{i+(j-1)M}^2$ is the realized volatility over the period $[(j-1)M, jM]$ for $j = 1, \ldots, \frac{1}{Mh}$. Note that when $q = 2$, $R_2 = RV$ (realized volatility). Similarly, for any

$q > 0$, define the integrated power variation by

$$\overline{\sigma^q} \equiv \int_0^1 \sigma_u^q du.$$

Mykland and Zhang (2009) show that $\frac{1}{c_{M,q}} R_q \overset{P}{\to} \overline{\sigma^q}$, where $c_{M,q} \equiv E\left(\left(\frac{\chi_M^2}{M}\right)^{q/2}\right)$ with χ_M^2 the standard χ^2 distribution with M degrees of freedom and

$$c_{M,q} = \left(\frac{2}{M}\right)^{q/2} \frac{\Gamma\left(\frac{q+M}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}, \tag{3.6}$$

where Γ is the Gamma function. Similarly, Mykland and Zhang (2009) provide a CLT result for R_q with M fixed, whereas Mykland and Zhang (2011) allow M to go to infinity with the sample size h^{-1}, provided M is of order $O(h^{-1/2})$. In particular, for $q = 2$, we have that under P and Q_h, as the number of intraday observations increases to infinity,

$$\frac{\sqrt{h^{-1}}\left(R_2 - \overline{\sigma^2}\right)}{\sqrt{V}} \overset{d}{\to} N(0,1), \tag{3.7}$$

where

$$V = \frac{M\left(c_{M,4} - c_{M,2}^2\right)}{c_{M,2}^2} \int_0^1 \sigma_u^4 du.$$

In practice, this result is infeasible since the asymptotic variance V depends on an unobserved quantity, the integrated quarticity $\int_0^1 \sigma_u^4 du$. Mykland and Zhang (2009) propose a consistent estimator of V ($\widehat{V} = \frac{M\left(c_{M,4} - c_{M,2}^2\right)}{c_{M,2}^2} \frac{1}{c_{M,4}} R_4$), and together with (3.7), we have the feasible CLT (cf. Remark 8 of Mykland and Zhang (2009)):

$$T_{h,M} \equiv \frac{\sqrt{h^{-1}}\left(R_2 - \overline{\sigma^2}\right)}{\sqrt{\hat{V}}} \overset{d}{\to} N(0,1).$$

Note that, when the block size $M = 1$, this result is equivalent to the CLT for realized volatility derived by Barndorff-Nielsen and Shephard (2002). In particular, $c_{1,2} = E\left(\chi_1^2\right) = 1$, and $c_{1,4} = E\left(\chi_1^2\right)^2 = 3$. Here, when $M > 1$, the realized volatility R_2 using the blocking approach is the same realized volatility studied by Barndorff-Nielsen and Shephard (2002), but the t-statistic is different because $\hat{V}$ changes with M. One advantage of the block-based estimator is to improve efficiency by varying the size of the block (see for e.g. Mykland, Shephard and Sheppard (2012)).

3.3.2 Bootstrap consistency

Here we show that the new bootstrap method we proposed in Section 2 is consistent when applied to realized volatility. Specifically, given (3.5) with $d = 1$, for $j = 1, \ldots, 1/Mh$, we let

$$y^*_{i+(j-1)M} = \sqrt{\frac{RV_{j,M}}{M}}\eta_{i+(j-1)M},\ j = 1, \ldots, M, \tag{3.8}$$

where $\eta_{i+(j-1)M} \sim$ i.i.d.$N(0,1)$ across (i,j). Note that this bootstrap method is related to the wild bootstrap approach proposed by Gonçalves and Meddahi (2009). In particular, when $M = 1$ and $d = 1$, it is equivalent to the wild bootstrap based on a standard normal external random variable.

We define the bootstrap realized volatility estimator as follows

$$R_2^* = \sum_{i=1}^{1/h} y_i^{*2} = \sum_{j=1}^{1/Mh} RV^*_{j,M},$$

where $RV^*_{j,M} = \sum_{i=1}^{M} y^{*2}_{i+(j-1)M}$. Letting

$$\frac{1}{M}\sum_{i=1}^{M}\eta^2_{i+(j-1)M} \equiv \frac{\chi^2_{j,M}}{M},$$

it follows that $RV^*_{j,M} = \frac{\chi^2_{j,M}}{M}RV_{j,M}$. We can easily show that

$$E^*\left(R_2^*\right) = c_{M,2}R_2,$$

and

$$V^* \equiv Var^*\left(h^{-1/2}R_2^*\right) = M\left(c_{M,4} - c^2_{M,2}\right)R_4.$$

Hence, we propose the following consistent estimator of V^*:

$$\hat{V}^* = M\frac{c_{M,4} - c^2_{M,2}}{c_{M,4}}R_4^*.$$

The bootstrap analogue of $T_{h,M}$ is given by

$$T^*_{h,M} \equiv \frac{\sqrt{h^{-1}}\left(R_2^* - c_{M,2}R_2\right)}{\sqrt{\hat{V}^*}}.$$

Theorem 3.3.1. *Suppose (3.1), (3.2) and (3.8) hold. If M is fixed or $M \to \infty$ as $h \to 0$*

such that $M = o\left(h^{-1/2}\right)$*, then as* $h \to 0$*,*

$$\sup_{x \in \Re} \left| P^* \left(T^*_{h,M} \leq x \right) - P \left(T_{h,M} \leq x \right) \right| \to 0,$$

in probability under Q_h *and under* P*.*

Theorem 3.3.1 provides a theoretical justification for using the bootstrap distribution of $T^*_{h,M}$ to estimate the distribution of $T_{h,M}$ under the general context studied by Mykland and Zhang (2009). This result also justifies the use of the bootstrap for constructing the studentized bootstrap (percentile-t) intervals.

Note that, when $M \to \infty$, such that $M = o(h^{-1/2})$, $V^* \xrightarrow{P} V$, we can also show that bootstrap percentile intervals for integrated volatility are valid. This is in contrast to the optimal two-point wild bootstrap proposed by Gonçalves and Meddahi (2009), which is only valid for percentile-t intervals.

3.4 Results for realized beta

3.4.1 Existing asymptotic theory and a new variance estimator

The goal of this section is to describe the realized beta in the context of Mykland and Zhang's (2009) blocking approach. In order to obtain a feasible CLT, we propose a consistent estimator of the variance of the realized beta, which is a new estimator in this literature. To derive this result, we use the approach of Dovonon, Gonçalves and Meddahi (2013) and suppose that σ is independent of W.[1] Note that contrary to Dovonon, Gonçalves and Meddahi (2013), we do not need here to suppose that $\mu_t = 0$ (since under Q_h high frequency returns have mean zero conditionally on σ).

For simplicity, we consider the bivariate case where $d = 2$ and look at results for assets k and l, whose ith high frequency returns in the jth block will be written as $y_{k,i+(j-1)M}$ and $y_{l,i+(j-1)M}$, respectively, for $i = 1, \ldots, M$ and $j = 1, \ldots, \frac{1}{Mh}$. It follows that under Q_h, $y_{l,i+(j-1)M} = \frac{1}{\sqrt{M}} C_{1(j)} \eta_{1,i+(j-1)M}$ and $y_{k,i+(j-1)M} = \frac{1}{\sqrt{M}} C_{21(j)} \eta_{1,i+(j-1)M} +$

[1] We make the assumption of no leverage for notational simplicity and because this allows us to easily compute the moments of the intraday returns conditionally on the volatility path. The same arguments would follow under the presence of leverage (for instance, by postulating a model for σ_t, as in Meddahi (2002)) but this would unnecessarily complicate the notation without any gain in the intuition.

$\frac{1}{\sqrt{M}} C_{2(j)} \eta_{2,i+(j-1)M}$, where

$$C_{(j)} \equiv \begin{pmatrix} C_{1(j)} & 0 \\ C_{21(j)} & C_{2(j)} \end{pmatrix} = \begin{pmatrix} \sqrt{\Gamma_{l(j)}} & 0 \\ \frac{\Gamma_{lk(j)}}{\sqrt{\Gamma_{l(j)}}} & \sqrt{\Gamma_{k(j)} - \frac{\Gamma^2_{lk(j)}}{\Gamma_{l(j)}}} \end{pmatrix},$$

$$\eta_{i+(j-1)M} \equiv \begin{pmatrix} \eta_{1,i+(j-1)M} \\ \eta_{2,i+(j-1)M} \end{pmatrix} \sim i.i.d.N(0, I_2),$$

I_2 is a 2×2 identity matrix, $\Gamma_{lk(j)} = \int_{(j-1)Mh}^{jMh} \Sigma_{lk}(u)\, du$, and when $k = l$, we write $\Gamma_{k(j)} = \Gamma_{kk(j)}$.

Then, conditionally on Σ, we can write

$$y_{li} = \beta_{lki} y_{ki} + u_i, \tag{3.9}$$

where independently across $i = 1, \ldots, 1/h$, $u_i | y_{ki} \sim N(0, V_i)$, with $V_i \equiv \Gamma_{li} - \frac{\Gamma^2_{lki}}{\Gamma_{ki}}$, and $\beta_{lki} \equiv \frac{\Gamma_{lki}}{\Gamma_{ki}}$, where $\Gamma_{lki} = \int_{(i-1)h}^{ih} \Sigma_{lk}(u)\, du$.

As Dovonon, Gonçalves and Meddahi (2013) argue, the conditional mean parameters of realized regression models are heterogeneous under stochastic volatility. This heterogeneity justifies why the pairs bootstrap method that they studied is not second-order accurate.

Under the approximation measure Q_h for the observables in the jth block ($j = 1, \ldots, \frac{1}{Mh}$), the regression (3.9) becomes

$$y_{l,i+(j-1)M} = \beta_{lk(j)} y_{k,i+(j-1)M} + u_{i+(j-1)M}, \tag{3.10}$$

where $u_{i+(j-1)M} | y_{k,i+(j-1)M} \sim$i.i.d.$N\left(0, V_{(j)}\right)$, for $i = 1, \ldots, M$, with $V_{(j)} \equiv \frac{1}{M}\left(\Gamma_{l(j)} - \frac{\Gamma^2_{lk(j)}}{\Gamma_{k(j)}}\right)$, and $\beta_{lk(j)} \equiv \frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}} = \frac{1}{Mh}\int_{(j-1)Mh}^{jMh} \beta_{lk}(u)\, du$. This implies that the integrated beta is $\beta_{lk} = Mh \sum_{j=1}^{1/Mh} \beta_{lk(j)} = \int_0^1 \beta_{lk}(u)\, du$.

Let us denote by $\hat{\beta}_{lk(j)}$ the ordinary least squares (OLS) estimator of $\beta_{lk(j)}$. Mykland and Zhang (2009) proposed to use $\hat{\beta}_{lk}$ defined as follows,

$$\hat{\beta}_{lk} = Mh \sum_{j=1}^{1/Mh} \hat{\beta}_{lk(j)} = Mh \sum_{j=1}^{1/Mh} \left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-1} \left(\sum_{i=1}^{M} y_{k,i+(j-1)M} y_{l,i+(j-1)M}\right),$$

to estimate the integrated beta. Note that the realized beta estimator studied by Dovonon, Gonçalves and Meddahi (2013) is a different statistic than ours. Here, the realized beta estimator $\hat{\beta}_{lk}$ is not directly a least squares estimator, but is the result of the average of

$\hat{\beta}_{lk(j)}$, the OLS estimators for each block. Since under Q_h, the volatility matrix is constant in each block j, we have that $\beta_{lki} = \beta_{lk(j)}$, for all $i = 1, \ldots, M$, implying consequently that the score is not heterogeneous and has mean zero. This simplifies the asymptotic inference on $\beta_{lk(j)}$, and on β_{lk}. Also note that contrary to what we have observed in the case of realized volatility estimator, here when $M = 1$, the realized beta estimator using the blocking approach become

$$\hat{\beta}_{lk} = h \sum_{i=1}^{1/h} \frac{y_{l,i}}{y_{k,i}},$$

which is a different statistic than the statistic studied by Barndorff-Nielsen and Shephard (2004). But when $M = h^{-1}$, both estimators are equivalent. However, as Mykland and Zhang (2011) pointed out, when $M \to \infty$ with the sample size h^{-1}, the local approximation is good only when $M = O(h^{-1/2})$. It follows then that we are not confortable to contrast Mykland and Zhang (2009) block-based realized beta estimator asymptotic results with those of Barndorff-Nielsen and Shephard (2004).

Mykland and Zhang (2009) provide a CLT result for β_{lk}. In particular, we have under P and Q_h, as the number of intraday observations increases to infinity (i.e. if $h \to 0$), by using Section 4.2 of Mykland and Zhang (2009),

$$\frac{\sqrt{h^{-1}}\left(\hat{\beta}_{lk} - \beta_{lk}\right)}{\sqrt{V_\beta}} \xrightarrow{d} N(0,1), \tag{3.11}$$

where

$$V_\beta = \begin{cases} \frac{M}{M-2} \int_0^1 \left(\frac{\Sigma_l(u)}{\Sigma_k(u)} - \beta_{lk}^2(u)\right) du, & \text{if } M = O(1), \text{ as } h \to 0 \text{ such that } M > 2\,(1+\delta) \text{ for any } \delta > 0, \\ \int_0^1 \left(\frac{\Sigma_l(u)}{\Sigma_k(u)} - \beta_{lk}^2(u)\right) du, & \text{if } M \to \infty \text{ as } h \to 0 \text{ such that } M = o(h^{-1/2}), \end{cases}$$

In practice, this result is infeasible since the asymptotic variance V_β depends on unobserved quantities. Mykland and Zhang (2009) did not provide any consistent estimator of V_β. One of our contributions is to propose a consistent estimator of V_β. To this end, we exploit the special structure of the regression model. To find the asymptotic variance of realized regression estimator $\hat{\beta}_{lk}$, we can write

$$\sqrt{h^{-1}}\left(\hat{\beta}_{lk} - \beta_{lk}\right) = M\sqrt{h} \sum_{j=1}^{1/Mh} \left(\hat{\beta}_{lk(j)} - \beta_{lk(j)}\right).$$

Since $\hat{\beta}_{lk(j)}$ are independent across j, it follows that

$$V_{\beta,h,M} \equiv Var\left(\sqrt{h^{-1}}\left(\hat{\beta}_{lk} - \beta_{lk}\right)\right) = M^2 h \sum_{j=1}^{1/Mh} Var\left(\hat{\beta}_{lk(j)} - \beta_{lk(j)}\right). \tag{3.12}$$

To compute (3.12), note that from standard regression theory, we have that under Q_h,

$$Var\left(\hat{\beta}_{lk(j)} - \beta_{lk(j)}\right) = E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\right) V_{(j)},$$

which implies that

$$V_{\beta,h,M} = M^2 h \sum_{j=1}^{1/Mh} E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\right) V_{(j)}. \tag{3.13}$$

Note that we can contrast V_β with equation (72) of Mykland and Zhang (2009). In fact, we can write under Q_h, $\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 \overset{d}{=} \frac{\Gamma_{k(j)}}{M} \sum_{i=1}^{M} v_{i+(j-1)M}^2 \overset{d}{=} \frac{\Gamma_{k(j)}}{M} \chi_{j,M}^2$, where $v_{i+(j-1)M} \sim i.i.d.N\left(0,1\right)$, and $\chi_{j,M}^2$ follow the standard χ^2 distribution with M degrees of freedom, and '$\overset{d}{=}$' denotes equivalence in distribution. Then for any integer $M > 2$ and conditionally on the volatility path, by using the expectation of the inverse of a Chi square distribution we have,

$$E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\right) = E\left(\frac{M}{\chi_{j,M}^2}\right) \Gamma_{k(j)}^{-1} = \frac{M}{M-2} \Gamma_{k(j)}^{-1}. \tag{3.14}$$

It follows then that

$$V_{\beta,h,M} = \frac{M}{M-2} \sum_{j=1}^{1/Mh} Mh \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right). \tag{3.15}$$

By using the structure of (3.13), a natural consistent estimator of $V_{\beta,h,M}$ is

$$\hat{V}_{\beta,h,M} \equiv M^2 h \sum_{j=1}^{1/Mh} \left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1} \left(\frac{1}{M-1} \sum_{i=1}^{M} \hat{u}_{i+(j-1)M}^2\right), \tag{3.16}$$

where $\hat{u}_{i+(j-1)M} = y_{l,i+(j-1)M} - \hat{\beta}_{lk(j)} y_{k,i+(j-1)M}$ (see Lemma 3.11.5 and Lemma 3.11.7 in the Appendix). Together with the CLT result (3.11), we have under P and Q_h the feasible result

$$T_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}}(\hat{\beta}_{lk} - \beta_{lk})}{\sqrt{\hat{V}_{\beta,h,M}}} \to^d N\left(0,1\right).$$

3.4.2 Bootstrap consistency

Here we show that the new bootstrap method we proposed in Section 2 is consistent when applied to realized betas. Specifically, given (3.5) with $d = 2$, for $j = 1, \ldots, 1/Mh$, we generate the M vector of returns as follows. For each $i = 1, \ldots, M$,

$$y^*_{i+(j-1)M} = \begin{pmatrix} y^*_{l,i+(j-1)M} \\ y^*_{k,i+(j-1)M} \end{pmatrix} = \frac{1}{\sqrt{M}} \begin{pmatrix} \sqrt{\hat{\Gamma}_{l(j)}}\eta_{1,i+(j-1)M} \\ \frac{\hat{\Gamma}_{lk(j)}}{\sqrt{\hat{\Gamma}_{l(j)}}}\eta_{1,i+(j-1)M} + \sqrt{\hat{\Gamma}_{k(j)} - \frac{\hat{\Gamma}^2_{lk(j)}}{\hat{\Gamma}_{l(j)}}}\eta_{2,i+(j-1)M} \end{pmatrix}, \tag{3.17}$$

where $\hat{\Gamma}_{l(j)} = \sum_{i=1}^{M} y^2_{l,i+(j-1)M}$, $\hat{\Gamma}_{k(j)} = \sum_{i=1}^{M} y^2_{k,i+(j-1)M}$, $\hat{\Gamma}_{lk(j)} = \sum_{i=1}^{M} y_{k,i+(j-1)M}y_{l,i+(j-1)M}$, and
$\begin{pmatrix} \eta_{1,i+(j-1)M} \\ \eta_{2,i+(j-1)M} \end{pmatrix} \sim i.i.d.N(0, I_2)$, I_2 is a 2×2 identity matrix.

Let $\hat{\beta}^*_{lk(j)}$ denote the OLS bootstrap estimator from the regression of $y^*_{l,i+(j-1)M}$ on $y^*_{k,i+(j-1)M}$ inside the block j. The bootstrap realized beta estimator is

$$\hat{\beta}^*_{lk} = Mh \sum_{j=1}^{1/Mh} \hat{\beta}^*_{lk(j)}.$$

It is easy to check that $\hat{\beta}^*_{lk}$ converges in probability (under P^*) to

$$\hat{\beta}_{lk} = Mh \sum_{j=1}^{1/Mh} E^* \left(\left(\sum_{i=1}^{M} y^{2*}_{k,i+(j-1)M} \right)^{-1} \left(\sum_{i=1}^{M} y^*_{k,i+(j-1)M} y^*_{l,i+(j-1)M} \right) \right).$$

The bootstrap analogue of the regression error $u_{i+(j-1)M}$ in model (3.10) is thus $u^*_{i+(j-1)M} = y^*_{l,i+(j-1)M} - \hat{\beta}_{lk(j)} y^*_{k,i+(j-1)M}$, whereas the bootstrap OLS residuals are defined as $\hat{u}^*_{i+(j-1)M} = y^*_{l,i+(j-1)M} - \hat{\beta}^*_{lk(j)} y^*_{k,i+(j-1)M}$. Thus, conditionally on the observed vector of returns $y_{i+(j-1)M}$, it follows that $u^*_{i+(j-1)M} | y^*_{k,i+(j-1)M} \sim i.i.d.N\left(0, \hat{V}_{(j)}\right)$, for $i = 1, \dots, M$, where

$$\hat{V}_{(j)} \equiv \frac{1}{M} \left(\hat{\Gamma}_{l(j)} - \frac{\hat{\Gamma}^2_{lk(j)}}{\hat{\Gamma}_{k(j)}} \right) = \frac{1}{M} \left(\sum_{i=1}^{M} y^2_{l,i+(j-1)M} - \frac{\left(\sum_{i=1}^{M} y_{k,i+(j-1)M} y_{l,i+(j-1)M} \right)^2}{\sum_{i=1}^{M} y^2_{k,i+(j-1)M}} \right).$$

We can show that

$$Var^* \left(\sqrt{h^{-1}} (\hat{\beta}^*_{lk} - \hat{\beta}_{lk}) \right) = \frac{M-1}{M-2} \hat{V}_{\beta,h,M}.$$

It follows then that a sufficient condition for the bootstrap to provide a consistent estimator of the asymptotic variance of $\sqrt{h^{-1}}(\hat{\beta}_{lk} - \beta_{lk})$ is to allow M to go to infinity. In particular when M increases with h^{-1} but at rate $o(h^{-1/2})$ (so that there is no contiguity adjustment), the bootstrap can be used to approximate the quantiles of the distribution of the root

$$\sqrt{h^{-1}} \left(\hat{\beta}_{lk} - \beta_{lk} \right),$$

thus justifying the construction of bootstrap percentile confidence intervals for β_{lk}. Our next theorem summarizes these results.

Theorem 3.4.1. *Consider DGP (3.1), (3.2) and suppose (3.17) holds. Then conditionally on σ, as $h \to 0$, under Q_h and P, the following hold*

a)

$$\begin{aligned} V^*_{\beta,h,M} &\equiv Var^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})\right) \\ &\overset{P}{\to} \begin{cases} \frac{M-1}{M-2} V_\beta, & \text{if } M = O(1), \text{ as } h \to 0 \text{ such that } M > 2\,(1+\delta) \text{ for any } \delta > 0, \\ V_\beta, & \text{if } M \to \infty \text{ as } h \to 0 \text{ such that } M = o(h^{-1/2}), \end{cases} \end{aligned}$$

b) $\sup_{x \in \mathbb{R}} \left| P^*\left(\sqrt{h^{-1}}\left(\hat{\beta}^*_{lk} - \hat{\beta}_{lk}\right) \le x\right) - P\left(\sqrt{h^{-1}}\left(\hat{\beta}_{lk} - \beta_{lk}\right)\right)\right| \overset{P}{\to} 0$, *as $h \to 0$ such that $M = o\left(h^{-1/2}\right)$.*

Part (a) of Theorem 3.4.1 shows that the bootstrap variance estimator is not consistent for V_β when the block size M is finite. But when the realized betas become an efficient estimator of integrated betas (i.e. if $M \to \infty$), we can use the bootstrap variance of $\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})$ to consistently estimate the covariance matrix V^*_β. Results in part (b) imply that the bootstrap realized beta estimator has a first order asymptotic normal distribution with mean zero and covariance matrix V_β. This is in line with the existing results in the cross section regression context, where the wild bootstrap and the pairs bootstrap variance estimator of the least squares estimator are robust to heteroskedasticity in the error term.

Bootstrap percentile intervals do not promise asymptotic refinements. Next, we propose a consistent bootstrap variance estimator that allows us to form bootstrap percentile-t intervals. More specifically, we can show that the following bootstrap variance estimator consistently estimates $V^*_{\beta,h,M}$:

$$\hat{V}^*_{\beta,h,M} \equiv M^2 h \sum_{j=1}^{1/Mh} \left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M}\right)^{-1} \left(\frac{1}{M-1}\sum_{i=1}^{M} \hat{u}^{*2}_{i+(j-1)M}\right). \tag{3.18}$$

Our proposal is to use this estimator to construct the bootstrap t-statistic, associated with the bootstrap realized regression coefficient $\hat{\beta}^*_{lk}$,

$$T^*_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\left(\hat{\beta}^*_{lk} - \hat{\beta}_{lk}\right)}{\sqrt{\hat{V}^*_{\beta,h,M}}}, \tag{3.19}$$

the bootstrap analogue of $T_{\beta,h,M}$.

Theorem 3.4.2. *Consider DGP (3.1), (3.2) and suppose (3.17) holds. Let $M > 4\,(2+\delta)$ for any $\delta > 0$ such that M is fixed or $M \to \infty$ as $h \to 0$ such that $M = o\left(h^{-1/2}\right)$, conditionally on σ, as $h \to 0$, the following hold.*

$$T^*_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\left(\hat{\beta}^*_{lk} - \hat{\beta}_{lk}\right)}{\sqrt{\hat{V}^*_{\beta,h,M}}} \to^{d^*} N(0,1)\,, \ \textit{in probability, under } Q_h \textit{ and } P.$$

Note that when the block size M is finite the bootstrap is also first order asymptotically valid when applied to the t-statistic $T^*_{\beta,h,M}$ (defined in (3.19)), as our Theorem 3.4.2 proves. This first order asymptotic validity occurs despite the fact that $V^*_{\beta,h,M}$ does not consistently estimate V_β when M is fixed. The key aspect is that we studentize the bootstrap OLS estimator with $\hat{V}^*_{\beta,h,M}$ (defined in (3.18)), a consistent estimator of $V^*_{\beta,h,M}$, implying that the asymptotic variance of the bootstrap t-statistic is one.

3.5 Higher-order properties

In this section, we investigate the asymptotic higher order properties of the bootstrap cumulants. Section 5.1 considers the case of realized volatility whereas Section 5.2 considers realized beta. The ability of the bootstrap to accurately match the cumulants of the statistic of interest is a first step to showing that the bootstrap offers an asymptotic refinement.

The results in this section are derived under the assumption of zero drift and no leverage (i.e. W is assumed independent of Σ). As in Dovonon, Gonçalves and Meddahi (2013), a nonzero drift changes the expressions of the cumulants derived here. The no leverage assumption is mathematically convenient as it allows us to condition on the path of volatility when computing the cumulants of our statistics. Allowing for leverage is a difficult but promising extension of the results derived here.

We introduce some notation. For any statistics T_h and T^*_h, we write $\kappa_j(T_h)$ to denote the j^{th} order cumulant of T_h and $\kappa^*_j(T^*_h)$ to denote the corresponding bootstrap cumulant. For $j = 1$ and 3, κ_j denotes the coefficient of the terms of order $O\left(\sqrt{h}\right)$ of the asymptotic expansion of $\kappa_j(T_h)$, whereas for $j = 2$ and 4, κ_j denotes the coefficients of the terms of order $O(h)$. The bootstrap coefficients $\kappa^*_{j,h}$ are defined similarly.

3.5.1 Higher order cumulants of realized volatility

Let $\sigma_{q,p} \equiv \frac{\overline{\sigma^q}}{\left(\overline{\sigma^p}\right)^{q/p}}$, for any $q, p > 0$, and $R_{q,p} \equiv \frac{R_q}{(R_p)^{q/p}}$. We make the following assumption.

Assumption H. The log price process follows (3.1) with $\mu_t = 0$ and σ_t is independent of W_t, where the volatility σ is a càdlàg process, bounded away from zero, and satisfies the following regularity condition:

$$\lim_{h\to 0} h^{(1/2)} \sum_{i=1}^{1/h} \left|\sigma_{\eta_i}^r - \sigma_{\xi_1}^r\right| = 0,$$

for some $r > 0$ and for any η_i and such that $0 \leq \xi_1 \leq \eta_1 \leq h \leq \xi_2 \leq \eta_2 \leq 2h \leq \ldots \leq \xi_{1/h} \leq \eta_{1/h} \leq 1$

Assumption H is stronger than required to prove the central limit theorem for R_q in Mykland and Zhang (2009), but it is a convenient assumption to derive the cumulants expansions of $T_{h,M}$ and $T_{h,M}^*$. Specifically, under Assumption H, Barndorff-Nielsen and Shephard (2004) show that for any $q > 0$, $\overline{\sigma_h^q} - \overline{\sigma^q} = o(\sqrt{h})$, where $\overline{\sigma_h^q} = h^{1-q/2} \sum_{s=1}^{1/h} \left(\int_{(s-1)h}^{sh} \sigma_u^2 du \right)^{q/2}$. Under (3.4) we have shown that for any positive integer $M \geq 1$, $\overline{\sigma_{h,M}^q} \equiv (Mh)^{1-q/2} \sum_{j=1}^{1/Mh} \left(\sigma_{j,M}^2\right)^{q/2} = \overline{\sigma_h^q}$ (see proof of Theorem 3.11.1 in Appendix F). It follows that under Q_h and Assumption H, $\overline{\sigma_{h,M}^q} - \overline{\sigma^q} = o(\sqrt{h})$ and similarly $R_q - c_{M,q}\overline{\sigma^q} = o_P(\sqrt{h})$, (this result also holds under Q_h), a result on which we subsequently rely on to establish the cumulants expansion of $T_{h,M}$ and $T_{h,M}^*$.

The following result states our main findings for realized volatility.

Proposition 3.5.1. *Consider DGP (3.1) and suppose (3.8) holds. Under Assumption H, conditionally on σ and under Q_h, and P, it follows that*

i) $\lim_{h\to 0} \kappa_{1,h,M}^* - \kappa_1 = \left(\frac{c_{M,6}}{\left(c_{M,4}\right)^{3/2}} - 1 \right) \left(-\frac{A_{1,M}}{2} \sigma_{6,4} \right)$, *which is nonzero if M is finite, and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.*

ii)

$$\lim_{h\to 0} \kappa_{2,h,M}^* - \kappa_2 = \left(\frac{c_{M,8}}{\left(c_{M,4}\right)^2} - 1 \right) \left(C_{1,M} - A_{2,M}\right) \sigma_{8,4} + \left(\frac{\left(c_{M,6}\right)^2}{\left(c_{M,4}\right)^3} - 1 \right) \left(\frac{7}{4} A_{1,M}^2 \sigma_{6,4}^2 \right),$$

which is nonzero if M is finite and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.

iii) $\lim_{h\to 0} \kappa_{3,h,M}^* - \kappa_3 = \left(\frac{c_{M,6}}{\left(c_{M,4}\right)^{3/2}} - 1 \right) \left(B_{1,M} - 3A_{1,M}\right) \sigma_{6,4}$, *which is nonzero if M is finite, and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.*

iv)

$$\begin{aligned} \lim_{h\to 0} \kappa_{4,h,M}^* - \kappa_4 &= \left(\frac{c_{M,6}}{\left(c_{M,4}\right)^{3/2}} - 1 \right) \left(B_{2,M} + 3C_{1,M} - 6A_{2,M}^2 \right) \sigma_{8,4} \\ &+ \left(\frac{\left(c_{M,6}\right)^2}{\left(c_{M,4}\right)^3} - 1 \right) \left(18{A_{1,M}}^2 - 6A_{1,M}B_{1,M} \right) \sigma_{6,4}^2 \end{aligned}$$

which is nonzero if M is finite and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.

Here, $A_{1,M}$, $A_{2,M}, C_{1,M}, C_{4,M}, C_{6,M}$, and $C_{8,M}$ are given in Lemma 3.11.2.

Proposition 3.5.1 shows that the cumulants of $T_{h,M}$ and $T^*_{h,M}$ do not agree when the block size M is fixed, implying that the bootstrap does not provide a higher-order asymptotic refinement for finite values of M. Nevertheless, when $M \to \infty$ at a rate $o(h^{-1/2})$ the bootstrap matches the first and third order cumulants through order $O\left(h^{-1/2}\right)$, which implies that it provides a second-order refinement, i.e. the bootstrap distribution $P^*\left(T^*_{h,M} \leq x\right)$ consistently estimates $P\left(T_{h,M} \leq x\right)$ with an error that vanishes as $o\left(h^{-1/2}\right)$ (assuming the corresponding Edgeworth expansions exist). This is in contrast with the first-order asymptotic Gaussian distribution whose error converges as $O\left(h^{-1/2}\right)$. Note that Gonçalves and Meddahi (2009) also proposed a choice of the external random variable for their wild bootstrap method which delivers second-order refinements. Our results for the bootstrap method based on the local Gaussianity are new. We will compare the two methods in the simulation section.

Parts (ii) and (iv) of Proposition 3.5.1 show that the new bootstrap method we propose is able to match the second and fourth order cumulants of $T_{h,M}$ when $M \to \infty$ as $h \to 0$ provided $M = o\left(h^{-1/2}\right)$. These results imply that the bootstrap distribution of $\left|T^*_{h,M}\right|$ consistently estimate the distribution of $\left|T_{h,M}\right|$ through order $O\left(h\right)$, in which case the bootstrap offers a third order asymptotic refinement (this again assumes that the corresponding Edgeworth expansions exist, something we have not attempted to prove in this chapter). If this is the case, then the local Gaussian bootstrap will deliver symmetric percentile$-t$ intervals for integrated volatility with coverage probabilities that converge to zero at the rate $o\left(h\right)$. In contrast, the coverage probability implied by the asymptotic theory-based intervals converge to the desired nominal level at the rate $O\left(h\right)$. The potential for the local Gaussian bootstrap intervals to yield third-order asymptotic refinements is particularly interesting because Gonçalves and Meddahi (2009) show that their wild bootstrap method is not able to deliver such refinements. Thus, our method is an improvement not only of the Gaussian asymptotic distribution but also of the best existing bootstrap methods.

Remark 3 One reason why the local Gaussian bootstrap is not able to match cumulants when M is finite is that the equation $\frac{c_{M,q}}{\left(c_{M,p}\right)^{q/p}} = 1$ does not always have an integer solution when $q, p \geq 1$. For instance, the equation $\frac{c_{M,6}}{\left(c_{M,4}\right)^{3/2}} = 1$ gives $M = -\frac{1}{4}$ as solution. However, we always have $\lim_{M\to\infty} \frac{c_{M,q}}{\left(c_{M,p}\right)^{q/p}} = 1$. This is the reason why the local Gaussian bootstrap is able to match cumulants when $M \to \infty$ (but not when M is finite).

3.5.2 Higher order cumulants of realized beta

In this section, we provide the first and third order cumulants of realized beta. These cumulants enter the Edgeworth expansions of the one-sided distribution functions of $T_{\beta,h,M}$ and $T^*_{\beta,h,M}$, $P^*\left(T^*_{\beta,h,M} \leq x\right)$ and $P\left(T_{\beta,h,M} \leq x\right)$, respectively.

Proposition 3.5.2. *Suppose (3.1), (3.2) and (3.17) hold. Conditionally on Σ, under Q_h and P, if M is fixed or $M \to \infty$ as $h \to 0$ such that $M = o\left(h^{-1/2}\right)$, then as $h \to 0$,*

i) $\lim_{h\to 0} \kappa^*_{1,\beta,h,M} - \kappa_{1,\beta} = 0$.

ii) $\lim_{h\to 0} \kappa^*_{3,\beta,h,M} - \kappa_{3,\beta} = 0$.

Proposition 3.5.2 shows that the cumulants of $T_{\beta,h,M}$ and $T^*_{\beta,h,M}$ agree through order $O\left(\sqrt{h}\right)$, which implies that the error of the bootstrap approximation $P^*\left(T^*_{\beta,h,M} \leq x\right)$ to the distribution of $T_{\beta,h,M}$ is of order $o\left(\sqrt{h}\right)$. Since the normal approximation has an error of the order $O\left(\sqrt{h}\right)$, this implies that the local Gaussian bootstrap is second-order correct. This result is an improvement over the bootstrap results in Dovonon, Gonçalves and Meddahi (2013), who showed that the pairs bootstrap is not second-order correct in the general case of stochastic volatility.

3.6 Monte Carlo results

In this section we assess by Monte Carlo simulation the accuracy of the feasible asymptotic theory approach of Mykland and Zhang (2009). We find that this approach leads to important coverage probability distortions when returns are not sampled too frequently. We also compare the finite sample performance of the new local Gaussian bootstrap method with the existing bootstrap method for realized volatility proposed by Gonçalves and Meddahi (2009).

For integrated volatility, we consider two data generating processes in our simulations. First, following Zhang et al. (2005), we use the one-factor stochastic volatility (SV1F) model of Heston (1993) as our data-generating process, i.e.

$$dX_t = \left(\mu - \nu_t/2\right) dt + \sigma_t dB_t,$$

and

$$d\nu_t = \kappa\left(\alpha - \nu_t\right) dt + \gamma\left(\nu_t\right)^{1/2} dW_t,$$

where $\nu_t = \sigma_t^2$, B and W are two Brownian motions, and we assume $Corr(B, W) = \rho$. The parameter values are all annualized. In particular, we let $\mu = 0.05/252$, $\kappa = 5/252$, $\alpha = 0.04/252$, $\gamma = 0.05/252$, $\rho = -0.5$.

We also consider the two-factor stochastic volatility (SV2F) model analyzed by Barndorff-Nielsen et al. (2009) and also by Gonçalves and Meddahi (2009), where [2]

$$
\begin{aligned}
dX_t &= \mu dt + \sigma_t dB_t, \\
\sigma_t &= s\text{-}exp\left(\beta_0, +\beta_1 \tau_{1t} + \beta_2 \tau_{2t}\right), \\
d\tau_{1t} &= \alpha_1 \tau_{1t} dt + dB_{1t}, \\
d\tau_{2t} &= \alpha_2 \tau_{2t} dt + \left(1 + \phi \tau_{2t}\right) dB_{2t}, \\
corr\left(dW_t, dB_{1t}\right) &= \varphi_1,\ corr\left(dW_t, dB_{2t}\right) = \varphi_2.
\end{aligned}
$$

We follow Huang and Tauchen (2005) and set $\mu = 0.03$, $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.00137$, $\alpha_2 = -1.386$, $\phi = 0.25$, $\varphi_1 = \varphi_2 = -0.3$. We initialize the two factors at the start of each interval by drawing the persistent factor from its unconditional distribution, $\tau_{10} \sim N\left(0, \frac{-1}{2\alpha_1}\right)$ and by starting the strongly mean-reverting factor at zero.

For integrated beta, the design of our Monte Carlo study follows that of Barndorff-Nielsen and Shephard (2004), and Dovonon Gonçalves and Meddahi (2013). In particular, we assume that $dX\left(t\right) = \sigma\left(t\right) dW\left(t\right)$, with $\sigma\left(t\right)\sigma'\left(t\right) = \Sigma\left(t\right)$, where

$$
\Sigma\left(t\right) = \begin{pmatrix} \Sigma_{11}\left(t\right) & \Sigma_{12}\left(t\right) \\ \Sigma_{21}\left(t\right) & \Sigma_{22}\left(t\right) \end{pmatrix} = \begin{pmatrix} \sigma_1^2\left(t\right) & \sigma_{12}\left(t\right) \\ \sigma_{21}\left(t\right) & \sigma_2^2\left(t\right) \end{pmatrix},
$$

and $\sigma_{12}\left(t\right) = \sigma_1\left(t\right)\sigma_2\left(t\right)\rho\left(t\right)$. As Barndorff-Nielsen and Shephard (2004), we let $\sigma_1^2\left(t\right) = \sigma_1^{2(1)}(t) + \sigma_1^{2(2)}(t)$, where for $s = 1, 2$, $d\sigma_1^{2(s)}(t) = -\lambda_s(\sigma_1^{2(s)}(t) - \xi_s)dt + \omega_s \sigma_1^{(s)}(t)\sqrt{\lambda_s} db_s(t)$, where b_i is the i-th component of a vector of standard Brownian motions, independent from W. We let $\lambda_1 = 0.0429$, $\xi_1 = 0.110$, $\omega_1 = 1.346$, $\lambda_2 = 3.74$, $\xi_2 = 0.398$, and $\omega_2 = 1.346$. Our model for $\sigma_2^2(t)$ is the GARCH(1,1) diffusion studied by Andersen and Bollerslev (1998): $d\sigma_2^2(t) = -0.035(\sigma_2^2(t) - 0.636)dt + 0.236\sigma_2^2(t)db_3(t)$. Finally, we follow Barndorff-Nielsen and Shephard (2004), and let $\rho(t) = (e^{2x(t)} - 1)/(e^{2x(t)} + 1)$, where x follows the GARCH diffusion: $dx(t) = -0.03(x(t) - 0.64)dt + 0.118x(t)db_4(t)$.

We simulate data for the unit interval $[0, 1]$. The observed log-price process X is generated using an Euler scheme. We then construct the h-horizon returns $y_i \equiv X_{ih} - X_{(i-1)h}$ based on samples of size $1/h$.

Tables 3.7 and 3.8 give the actual rates of 95% confidence intervals of integrated volatility and integrated beta, computed over 10,000 replications. Results are presented for six different samples sizes: $1/h$ 1152, 576, 288, 96, 48, and 12, corresponding to "1.25-minute", "2.5-minute", "5-minute", "15-minute", "half-hour" and "2-hour" returns. In

[2] The function s-exp is the usual exponential function with a linear growth function splined in at high values of its argument: $s\text{-}exp(x) = exp(x)$ if $x \leq x_0$ and $s\text{-}exp(x) = \frac{exp(x_0)}{\sqrt{x_0 - x_0^2 + x^2}}$ if $x > x_o$, with $x_0 = log(1.5)$.

Table 3.7, for each sample size we have computed the coverage rate by varying the block size, whereas in Table 3.8 we summarize results by selecting the optimal block size. We also report results for confidence intervals based on a logarithmic version of the statistic $T_{h,M}$ and its bootstrap version.

In our simulations, bootstrap intervals use 999 bootstrap replications for each of the 10,000 Monte Carlo replications. We consider the studentized (percentile-t) symmetric bootstrap confidence interval method computed at the 95% level.

As for all blocking methods, to implement our bootstrap methods, we need to choose the block size M. We follow Politis and Romano (1999) and Hounyo, Gonçalves and Meddahi (2013) and use the Minimum Volatility Method. Here we describe the algorithm we employ for a two-sided confidence interval.

Algorithm: Choice of the block size M by minimizing confidence interval volatility

(i) For $M = M_{small}$ to $M = M_{big}$ compute a bootstrap interval for the parameter of interest (integrated volatility or integrated beta) at the desired confidence level, this resulting in endpoints $IC_{M,low}$ and $IC_{M,up}$.

(ii) For each M compute the volatility index VI_M as the standard deviation of the interval endpoints in a neighborhood of M. More specifically, for a smaller integer l, let VI_M equal to the standard deviation of the endpoints $\{IC_{M-l,low}, \ldots, IC_{M+l,low}\}$ plus the standard deviation of the endpoints $\{IC_{M-l,up}, \ldots, IC_{M+l,up}\}$, i.e.

$$VI_M \equiv \sqrt{\frac{1}{2l+1}\sum_{i=-l}^{l}\left(IC_{M+i,low} - \bar{IC}_{low}\right)^2} + \sqrt{\frac{1}{2l+1}\sum_{i=-l}^{l}\left(IC_{M+i,up} - \bar{IC}_{up}\right)^2},$$

where $\bar{IC}_{low} = \frac{1}{2l+1}\sum_{i=-l}^{l} IC_{M+i,low}$ and $\bar{IC}_{up} = \frac{1}{2l+1}\sum_{i=-l}^{l} IC_{M+i,up}$.

(iii) Pick the value M^* corresponding to the smallest volatility index and report $\{IC_{M^*,low}, IC_{M^*,up}\}$ as the final confidence interval.

One might ask what is a selection of reasonable M_{small} and M_{big}? In our experience, for a sample size $1/h = 1152$, the choices $M_{small} = 1$ and $M_{big} = 12$ usually suffice, for the samples sizes : $1/h = 1152$, 576, 288, 96, and 48, we have used $M_{small} = 1$ and $M_{big} = 12$. For results in Table 3.8, we used $l = 2$ in our simulations. Some initial simulations (not recorded here) showed that the actual coverage rate of the confidence intervals using the bootstrap is not sensitive to reasonable choice of l, in particular, for $l = 1, 2, 3$.

Starting with integrated volatility, the Monte Carlo results in Tables 3.7 and 3.8 show that for both models (SV1F and SV2F), the asymptotic intervals tend to undercover.

The degree of undercoverage is especially large, when sampling is not too frequent. It is also larger for the raw statistics than for the log-based statistics. The SV2F model exhibits overall larger coverage distortions than the SV1F model, for all sample sizes. When $M = 1$, the Gaussian bootstrap method is equivalent to the wild bootstrap of Gonçalves and Meddahi (2009) that uses the normal distribution as external random variable. One can see that the bootstrap replicates their simulations results. In particular, the Gaussian bootstrap intervals tend to overcover across all models. The actual coverage probabilities of the confidence intervals using the Gaussian bootstrap are typically monotonically decreasing in M, and does not tend to decrease very fast in M for larger values of sample size.

A comparison of the local Gaussian bootstrap with the best existing bootstrap methods for realized volatility[3] shows that, for smaller samples sizes, the confidence intervals based on Gaussian bootstrap are conservative, yielding coverage rates larger than 95% for the SV1F model. The confidence intervals tend to be closer to the desired nominal level for the SV2F than the best bootstrap proposed by Gonçalves and Meddahi (2009). For instance, for SV1F model, the Gaussian bootstrap covers 96.51% of the time when $h^{-1} = 12$ whereas the best bootstrap of Gonçalves and Meddahi (2009) does only 87.42%. These rates decrease to 93.21% and 80.42% for the SV2F model, respectively.

We also consider intervals based on the i.i.d. bootstrap studied by Gonçalves and Meddahi (2009). Despite the fact that the i.i.d. bootstrap does not theoretically provide an asymptotic refinement for two-sided symmetric confidence intervals, it performs well.

While none of the intervals discussed here (bootstrap or asymptotic theory-based) allow for $M = h^{-1}$, we have also studied this setup which is nevertheless an obvious interest in practice. For the SV1F model, results are not very sensitive to the choice of the block size, whereas for the SV2F model coverage rates for intervals using a very large value of block size ($M = h^{-1}$) are systematically much lower than 95% even for the largest sample sizes. When $M = h^{-1}$, the realized volatility R_2 using the blocking approach is the same realized volatility studied by Barndorff-Nielsen and Shephard (2002), but the estimator of integrated quarticity using the blocking approach is $\frac{h^{-1}+2}{h^{-1}} R_2^2$. This means that asymptotically we replace $\int_0^1 \sigma_t^4 dt$ by $\left(\int_0^1 \sigma_t^2 dt\right)^2$, which is only valid under constant volatility. By Cauchy-Schwarz inequality, we have $\left(\int_0^1 \sigma_t^2 dt\right)^2 \leq \int_0^1 \sigma_t^4 dt$, it follows then that we underestimated the asymptotic variance of the realized volatility estimator. This explains the poor performance of the theory based on the blocking approach when the block size is too large. This also confirms the theoretical prediction, which require $M = O(\sqrt{h^{-1}})$ for a good approximation for the probability measure P.

[3]The wild bootstrap based on Proposition 4.5 of Gonçalves and Meddahi (2009).

For realized beta, we see that intervals based on the feasible asymptotic procedure using Mykland and Zhang's (2009) blocking approach and the bootstrap tend to be similar for larger sample sizes whereas, at the smaller sample sizes, intervals based on the asymptotic normal distribution are quite severely distorted. For instance, the coverage rate for the feasible asymptotic theory of Mykland and Zhang (2009) when $h^{-1} = 12$ (cf. $h^{-1} = 48$) is only equal to 88.49% (92.86%), whereas it is equal to 95.17% (94.84%), for the Gaussian bootstrap (the corresponding symmetric interval based on the pairs bootstrap of Dovonon Gonçalves and Meddahi (2013) yields a coverage rate of 93.59% (93.96%), better than Mykland and Zhang (2009) but worse than the Gaussian bootstrap interval). Our Monte Carlo results also confirm that for a good approximation, a very large block size is not recommended.

Overall, all methods behave similarly for larger sample sizes, in particular the coverage rate tends to be closer to the desired nominal level. The Gaussian bootstrap performance is quite remarkable and outperforms the existing methods, especially for smaller samples sizes ($h^{-1} = 12$ and 48).

3.7 Empirical results

As a brief illustration, in this section we implement the local Gaussian bootstrap method with real high-frequency financial intraday data, and compare it to the existing feasible asymptotic procedure of Mykland and Zhang (2009). The data consists of transaction log prices of General Electric (GE) shares carried out on the New York Stock Exchange (NYSE) in August 2011. Before analyzing the data we have cleaned the data. For each day, we consider data from the regular exchange opening hours from time stamped between 9:30 a.m. till 4 p.m. Our procedure for cleaning data is exactly identical to that used by Barndorff-Nielsen et al. (2008). We detail in Appendix E the cleaning we carried out on the data.

We implemented the realized volatility estimator of Mykland and Zhang (2009) on returns recorded every S transactions, where S is selected each day so that there are 96 observations a day. This means that on average these returns are recorded roughly every 15 minutes. Table 3.9 in the Appendix provides the number of transactions per day, and the sample size used. Typically each interval corresponds to about 131 transactions.

This choice is motivated by the empirical study of Hansen and Lunde (2006), who investigate 30 stocks of the Dow Jones Industrial Average, in particular they have presented detailed work for the GE shares. They suggest to use 10 to 15 minutes horizon for liquid assets to avoid the market microstructure noise effect.

Hence the main assumptions underlying the validity of the Mykland and Zhang (2009) block-based method and our new bootstrap method are roughly satisfied and we feel comfortable to implement them on this data.

To implement the realized volatility estimator, we need to choose the block size M. We use the Minimum Volatility Method described above to choose M.

We consider bootstrap percentile-t intervals, computed at the 95% level. The results are displayed in Figure 3.3 in the appendix E in terms of daily 95% confidence intervals (CIs) for integrated volatility. Two types of intervals are presented: our proposed new local Gaussian bootstrap method , and the the feasible asymptotic theory using Mykland and Zhang (2009) blocking approach. The realized volatility estimate R_2 is in the center of both confidence intervals by construction. A comparison of the local Gaussian bootstrap intervals with the intervals based on the feasible asymptotic theory using Mykland and Zhang (2009) block-based approach suggests that the both types of intervals tend to be similar. The width of these intervals varies through time. However there are instances where the bootstrap intervals are wider than the asymptotic theory-based interval. These days often correspond to days with large estimate of volatility. We have asked whether it will be due to jumps. At this end we have implemented the jumps test using blocked bipower variation of Mykland, Shephard and Sheppard (2012). We have found no evidence of jumps at 5% significance level for these two days. The figures also show a lot of variability in the daily estimate of integrated volatility.

3.8 Conclusion

This chapter proposes a new bootstrap method for statistics that are smooth functions of the realized multivariate volatility matrix based on Mykland and Zhang's (2009) blocking approach. We show how and to what extent the local Gaussianity assumption can be explored to generate a bootstrap approximation. We use Monte Carlo simulations and derive higher order expansions for cumulants to compare the accuracy of the bootstrap and the normal approximations at estimating confidence intervals for integrated volatility and integrated beta. Based on these expansions, we show that at second order the bootstrap matches the cumulants of realized betas-based t-statistics whereas it provides a third-order asymptotic refinement for realized volatility. This is an improvement of the existing bootstrap results. Our new bootstrap method also generalizes the wild bootstrap of Gonçalves and Meddahi (2009). Monte Carlo simulations suggest that the Gaussian bootstrap improves upon the first-order asymptotic theory in finite samples and outperform the existing bootstrap methods for realized volatility and realized betas. An important extension is to prove the validity of the Edgeworth expansions derived here. Another promising extension is to use the bootstrap method for volatility estimator (multipower variation) using the blocking approach in presence of jumps.

Conclusion Générale

A travers trois chapitres, cette thèse propose différentes méthodes de bootstrap pour faire l'inférence sur la volatilité intégrée ou des mesures de Co-volatilité comme les betas. Ces méthodes ont été développées essentiellement dans deux contextes théoriques. Dans un premier temps nous considérons un cadre théorique où les prix des actifs financiers ont été contaminés par le bruit de microstructure. Le second cadre théorique est celui qui se base sur l'hypothèse de Gaussianité locale des données financières de haute fréquence, en particulier l'approche par bloc proposée par Mykland et Zhang (2009). Nous montrons la validité théorique des méthodes de bootstrap. Nous montrons également en utilisant des expansions d'Edgeworth et des simulations Monte Carlo, que grâce aux nouvelles méthodes de bootstrap proposées dans cette thèse, les distributions des statistiques d'intérêt sont mieux estimées, comparativement à la théorie asymptotique et aux méthodes de bootstrap existantes. Enfin nous illustrons toutes les méthodes de bootstrap que nous avions développées dans cette thèse en utilisant des données financières réelles. Notre agenda de recherche contient plusieurs pistes de travail. Nous prévoyons développer des expansions d'Edgeworth pour les estimateurs de volatilité qui utilisent la méthode de "pré-moyennement". Nous envisageons aussi justifier la validité théorique de ces expansions. Etudiez la méthode de bootstrap développée dans le Chapitre 2, dans un contexte multi-varié. Enfin montrer la validité du bootstrap en présence des sauts dans les prix des actifs.

Bibliography

[1] Andersen, T.G. and T. Bollerslev (1998). "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885-905.

[2] Bandi, F., and J. Russell (2008). "Microstructure noise, realized variance, and optimal sampling,.Review of Economic Studies," 75(2), 339.369.

[3] Bandi, F., and J. Russell (2011). "Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations", *Journal of Econometrics*, 160, 145-159.

[4] Barndorff-Nielsen, O., S.E. Graversen, J. Jacod, and N. Shephard (2006). "Limit theoreems for bipower variation in financial econometrics," *Econometric Theory*, 22, 677-719.

[5] Barndorff-Nielsen, O., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). "A central limit theorem for realised power and bipower variations of continuous semimartingales." In Y. Kabanov, R. Lipster, and J. Stoyanov (Eds.), From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev, 33 - 68. Springer.

[6] Barndorff-Nielsen, O. and N. Shephard (2002). "Econometric analysis of realized volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society*, Series B, 64, 253-280.

[7] Barndorff-Nielsen, O. and N. Shephard (2004). "Power and bipower variation with stochastic volatility and jumps," *Journal of Financial Econometrics*, 2, 1-48.

[8] Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2008). "Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise," *Econometrica*,76, 1481-1536

[9] Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2008b). "Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading," Working paper, Oxford University.

[10] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., (2009). "Realised kernels in practice: trades and quotes, " *Econometrics Journal* 12, C1-C32.

[11] Bühlmann, P. and H. R. Künsch (1995): "The blockwise bootstrap for general parameters of a stationary time series," *Scandinavian Journal of Statistics*, 22(1), 35-54.

[12] Campbell, J., A. Sunderam, and L. Viceira (2008). "Inflation bets or deflation hedges? The changing risks of nominal bonds," Working paper, Harvard University.

[13] Christensen, K., S. Kinnebrock, and M. Podolskij (2010). "Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data," *Journal of Econometrics*, 159, 116-133.

[14] Christensen, K., R. Oomen and M. Podolskij (2010). quantile-based estimation of the integrated variance," *Journal of Econometrics 159*, 74-98.

[15] Diebold, F.X. and Strasser, G.H. (2012). "On the correlation structure of microstructure noise: a financial economic approach", *Review of Economics Studies*, forthcoming.

[16] French, K.R., G.W. Schwert and R.F. Stambaugh (1987). "Expected stock returns and volatility," *Journal of Financial Economics*, 19, 3-29.

[17] Dovonon, P., Gonçalves, S. and N. Meddahi (2013). " Bootstrapping realized multivariate volatility measures," *Journal of Econometrics* 172,49-65.

[18] Duffie D., (1996). " Dynamic asset pricing theory," Princeton, NJ: Princeton University Press.

[19] Ghysels, E.,A.C. Harvey, and E. Renault, (1996). Stochastic volatility in Handbook of Statistics: Volume 14, ed. by G.S. Maddala, and C.R. Rao. North-Holland, pp. 119-191.

[20] Gobbi, F. and C. Mancini, (2006). "Identifying the diffusion covariation and the co-jumps given discrete observations", Working paper, University of Firenze.

[21] Gonçalves, S. and N. Meddahi (2009). "Bootstrapping realized volatility," *Econometrica*, 77(1), 283.306.

[22] Gonçalves, S. and N. Meddahi (2008). "Edgeworth Corrections for Realized Volatility", *Econometric Reviews*, 27 (1), 139-162.

[23] Gonçalves, S., Hounyo, U. and N. Meddahi (2013). "Bootstrap inference for pre-averaged realized volatility based on non-overlapping returns", manuscript.

[24] Gonçalves, S., and H. White (2002). "The bootstrap of the mean for dependent heterogeneous arrays," *Econometric Theory*, 18, 1367-1384.

[25] Griffin, J., and R. Oomen, (2006). "Covariance measurement in the presence of non-synchronous trading and market microstructure noise," working paper, Warwick University.

[26] Hall, P., (1992). The bootstrap and Edgeworth expansion. Springer-Verlag, New York.

[27] Hansen, P.R. and A. Lunde (2006). "Realized variance and market microstructure noise,"*Journal of Business and Economics Statistics*, 24, 127-161.

[28] Hautsch N., and Podolskij, M., (2012). "Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, Implementation, and Empirical Evidence," *Forthcoming Journal of Business and Economic Statistics.*

[29] Heston, S. (1993). "Closed-form solution for options with stochastic volatility with applications to bonds and currency options," *Review of Financial Studies, 6, 327-343.*

[30] Hounyo, U. , Gonçalves, S., and N. Meddahi (2013). "Bootstrapping pre-averaged realized volatility under market microstructure noise," manuscript.

[31] Jacod, J., (1994). "Limit of random measures associated with the increments of a Brownian semimartingale," Preprint number 120, Laboratoire de Probabilitités, Université Pierre et Marie Curie, Paris.

[32] Jacod, J. and P. Protter (1998). "Asymptotic error distributions for the Euler method for stochastic differential equations," *Annals of Probability* 26, 267-307.

[33] Jacod, J., Y. Li, P. Mykland, M. Podolskij, and M. Vetter (2009). "Microstructure noise in the continuous case: the pre-averaging approach," Stochastic Processes and Their Applications.

[34] Kalnina, I., and O. Linton (2008). "Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error," Journal of Econometrics, 147(1), 47.59.

[35] Kalnina, I., (2011). "Nonparametric tests of time variation in betas," " manuscript.

[36] Katz, M.L., (1963). "Note on the Berry-Esseen theorem," *Annals of Mathematical Statistics* 34, 1107-1108.

[37] Künsch, H.R. (1989). "The jackknife and the bootstrap for general stationary observations," *Annals of Statistics* 17, 1217-1241.

[38] Podolskij, M., Ziggel, D., (2007). "Boostrapping bipower variation," Technical Report. Ruhr-University of Bochum.

[39] Podolskij, M., and M. Vetter (2009). "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps," *Bernoulli*, 15(3), 634.658.

[40] Politis, D. N. and Romano, J. P. (1992). "A general resampling scheme for triangular arrays of α-mixing random variables," *Annals of Statistics* 20, 1985-2007.

[41] Politis, D.N., Romano, J.P., Wolf, M., (1999). "Subsampling," Springer-Verlag, New York.

[42] Liu, R.Y. (1988). "Bootstrap procedure under some non-i.i.d. models," *Annals of Statistics* 16, 1696-1708.

[43] Mammen, E., (1993). "Bootstrap and wild bootstrap for high dimensional linear models," *Annals of Statistics* 21, 255-285.

[44] Meddahi, N., (2002). "A theoretical comparison between integrated and realized volatility," *Journal of Applied Econometrics* 17, 475-508.

[45] Mykland, P. and L. Zhang (2009). "Inference for continous semimartingales observed at high frequency," *Econometrica* 77, 1403-1455.

[46] Mykland, P.A. and L. Zhang (2009). "Inference for continous semimartingales observed at high frequency," *Econometrica* 77, 1403-1455.

[47] Mykland, P.A. and L. Zhang (2011). "The double Gaussian approximation for high frequency data," *Scandinavian Journal of Statistics* 38, 215-236.

[48] Mykland, P.A. N. Shephard, and K. Shepphard (2012). "Efficient and feasible inference for the components of financial variation using blocked multipower variation," Working paper, Oxford University.

[49] Ross, S.M., (1976). "The arbitrage theory of capital asset pricing," *Journal of Econometrics Theory* 13, 341-360.

[50] Vetter, M. (2008). "Estimation methods in noisy diffusion models," *Ph.D. Thesis. Bochum University.*

[51] Viceira, L.M., (2007). "Bond Risk, Bond Return Volatility, and the Term Structure of Interest Rates," working paper, Harvard Business School.

[52] Wu, C.F.J., (1986). "Jackknife, bootstrap and other resampling methods in regression analysis," *Annals of Statistics 14*, 1261-1295.

[53] Xiu, D. (2010). “Quasi-maximum likelihood estimation of volatility with high frequency data,” *Journal of Econometrics*, 159, 235–250.

[54] Zhang, L. (2011). “Estimating covariation: Epps effect, microstructure noise,” *Journal of Econometrics*, 160, 33-47.

[55] Zhang, L, P.A. Mykland, and Y. Aït-Sahalia (2005). “A tale of two time-scales: determining integrated volatility with noisy high frequency data,” *Journal of the American Statistical Association*, 100, 1394-1411.

[56] Zhang, L., Mykland, P. and Y. Aït-Sahalia (2011). “Edgeworth expansions for realized volatility and related estimators,”*Journal of Econometrics*, 166, 213-223.

[57] Zhang, L., Mykland, P. and Y. Aït-Sahalia (2011a). “Edgeworth expansions for realized volatility and related estimators,” *Journal of Econometrics*, 160, 190-203.

[58] Zhang, L., Mykland, P. and Y. Aït-Sahalia (2011b). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160, 160-165.

Appendices

3.9 Appendix for Chapter 1

3.9.1 Appendix A

Table 3.1 reports the actual coverage rates for the feasible asymptotic theory approach and for our bootstrap methods. Table 3.2 contains the finite sample values of the first four cumulants of T_n and its bootstrap analogue T_n^*. In Table 3.3 we provide some statistics of GE shares in December 2011.

Table 3.1. Coverage rate of Nominal 95 % intervals

	SV1F							SV2F						
		WB1		WB2		WB3			WB1		WB2		WB3	
n	CLT	Perc	Perc-t	Perc	Perc-t	Perc	Perc-t	CLT	Perc	Perc-t	Perc	Perc-t	Perc	Perc-t
$\xi^2 = 0.0001$														
195	77.54	77.49	97.91	76.42	91.05	61.11	81.41	69.49	69.38	94.72	68.51	86.78	55.51	71.89
390	84.85	84.47	98.42	83.51	93.71	66.76	90.20	77.97	77.64	96.17	76.89	89.88	62.87	82.42
780	86.82	86.11	98.43	85.41	93.94	67.73	92.91	80.61	80.19	96.24	79.09	90.17	63.36	85.87
1560	88.89	88.13	98.36	87.74	93.93	69.58	94.36	83.36	82.89	96.63	82.03	90.87	65.16	89.07
4680	91.49	90.65	98.63	90.78	94.69	72.56	96.59	86.17	85.59	96.76	85.41	91.74	67.71	91.92
7800	92.78	92.04	98.56	92.34	95.12	73.24	96.97	89.50	88.66	97.46	88.59	93.33	70.11	94.21
11700	93.01	92.41	98.35	92.63	95.11	73.40	97.16	89.05	88.34	97.09	88.27	93.15	70.23	94.15
23400	93.48	92.85	98.06	93.09	94.89	74.26	97.56	89.89	89.06	96.86	89.33	92.81	71.13	94.67
$\xi^2 = 0.001$														
195	77.63	77.46	97.90	76.56	90.92	61.18	81.54	69.72	69.81	94.78	68.83	86.59	55.86	71.86
390	85.02	84.48	98.50	83.66	93.75	66.71	90.38	77.95	77.73	96.14	76.97	89.93	63.17	82.57
780	86.81	86.11	98.43	85.22	93.86	67.91	92.76	80.55	80.23	96.14	79.36	90.25	63.68	85.94
1560	88.91	88.13	98.48	87.74	93.94	69.51	94.46	83.26	82.70	96.66	82.10	90.91	65.16	89.12
4680	91.47	90.76	98.67	90.78	94.78	72.55	96.59	86.33	85.66	96.68	85.37	91.94	67.86	91.93
7800	92.86	91.91	98.56	92.37	95.06	73.47	97.00	89.56	88.73	97.54	88.53	93.36	70.09	94.25
11700	92.98	92.25	98.32	92.56	95.14	73.57	97.10	88.95	88.19	97.00	88.15	93.04	70.16	94.23
23400	93.45	92.88	98.12	93.12	94.92	74.18	97.51	90.01	89.16	96.80	89.40	92.83	71.22	94.80
$\xi^2 = 0.01$														
195	77.93	77.67	97.67	76.86	91.23	61.69	81.35	70.17	70.12	94.80	69.28	86.74	56.07	72.73
390	85.09	84.57	98.35	83.61	93.59	66.85	90.28	78.59	78.46	96.42	77.43	90.00	62.89	83.60
780	86.75	86.29	98.38	85.31	93.47	67.96	92.75	81.29	80.90	96.33	79.92	90.36	63.86	86.54
1560	89.03	88.12	98.41	87.74	94.02	69.16	94.54	83.45	82.68	96.51	82.20	91.06	65.40	89.59
4680	91.42	90.54	98.78	90.66	94.62	72.39	96.64	86.78	86.04	96.57	85.67	91.97	68.07	92.17
7800	92.61	91.77	98.63	92.24	94.90	73.48	97.03	89.41	88.67	97.50	88.65	93.26	70.26	94.31
11700	93.22	92.36	98.43	92.72	94.92	73.63	97.17	89.09	88.4	96.97	88.42	92.93	70.17	94.30
23400	93.40	92.89	98.09	93.10	94.75	74.20	97.58	90.13	89.33	96.79	89.41	92.96	71.17	94.71

Notes: CLT-intervals based on the Normal; WB1 wild bootstrap intervals based on the external random variable WB1; WB2 wild bootstrap intervals based on the external random variable WB2; WB3 wild bootstrap intervals based on the external random variable WB3. 10,000 Monte Carlo trials with 999 bootstrap replications each.

TABLE 3.2. Summary results for the stutentized statistic T_n and its bootstrap analogue $T_n^{\star}$

	SV1F				SV2F			
$\xi^2 = 0.01$	T_n	T_n^{*WB1}	T_n^{*WB2}	T_n^{*WB3}	T_n	T_n^{*WB1}	T_n^{*WB2}	T_n^{*WB3}
$n = 195$								
Mean	-1.109	-1.802	-0.552	-0.413	-1.798	-2.319	-0.752	-0.440
Standard error	2.356	4.526	2.135	1.298	3.921	5.867	2.466	1.264
Excess Skewness	-3.717	-5.751	-3.024	-0.009	-6.000	-6.041	-2.724	0.158
Excess Kurtosis	29.279	65.763	14.765	-1.206	72.426	71.078	11.29	-1.255
Cov two-sided	77.93	97.67	91.23	81.35	70.17	94.8	86.74	72.73
$n = 390$								
Mean	-0.594	-1.275	-0.393	-0.373	-0.997	-1.676	-0.590	-0.410
Standard error	1.661	2.969	1.755	1.317	2.325	3.776	2.119	1.283
Excess Skewness	-2.454	-3.940	-2.685	-0.169	-3.235	-4.218	-2.591	0.022
Excess Kurtosis	11.775	31.511	12.71	-0.991	18.615	35.287	10.775	-1.13
Cov two-sided	85.09	98.35	93.59	90.28	78.59	96.42	90.00	83.60
$n = 780$								
Mean	-0.519	-0.991	-0.297	-0.338	-0.850	-1.342	-0.482	-0.384
Standard error	1.464	2.304	1.517	1.323	1.982	2.924	1.884	1.293
Excess Skewness	-1.973	-2.974	-2.235	-0.279	-2.620	-3.306	-2.391	-0.077
Excess Kurtosis	8.798	17.221	9.053	-0.767	12.026	21.371	9.370	-1.005
Cov two-sided	86.75	98.38	93.47	92.75	81.29	96.33	90.36	86.54
$n = 1560$								
Mean	-0.409	-0.788	-0.228	-0.299	-0.669	-1.094	-0.395	-0.357
Standard error	1.297	1.909	1.353	1.323	1.704	2.389	1.690	1.300
Excess Skewness	-1.369	-2.381	-1.829	-0.366	-2.203	-2.672	-2.144	-0.166
Excess Kurtosis	3.714	10.773	6.193	-0.519	9.133	13.287	7.579	-0.858
Cov two-sided	89.03	98.41	94.02	94.54	83.45	96.51	91.06	89.59
$n = 23400$								
Mean	-0.191	-0.375	-0.095	-0.170	-0.334	-0.560	-0.183	-0.247
Standard error	1.064	1.277	1.085	1.281	1.225	1.490	1.247	1.295
Excess Skewness	-0.592	-1.226	-0.859	-0.419	-1.117	-1.504	-1.331	-0.364
Excess Kurtosis	0.599	2.611	1.357	0.066	2.518	3.794	3.053	-0.298
Cov two-sided	93.40	98.09	94.75	97.58	90.13	96.79	92.96	94.71

Notes: T_n studentized statistic; T_n^{*WB1} studentized wild bootstrap statistic based on WB1; T_n^{*WB2} studentized wild bootstrap statistic based on WB2; T_n^{*WB3} studentized wild bootstrap statistic based on WB3. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 3.3. Summary statistics

Days	Trans	n	S
1 Dec	11924	1491	8
2 Dec	11681	1461	8
5 Dec	10538	1506	7
6 Dec	12959	1440	9
7 Dec	11360	1420	8
8 Dec	10064	1438	7
9 Dec	12120	1515	8
12 Dec	12082	1511	8
13 Dec	10379	1483	7
14 Dec	12616	1577	8
15 Dec	10869	1553	7
16 Dec	12265	1534	8
19 Dec	11119	1589	7
20 Dec	12623	1578	8
21 Dec	13270	1475	9
22 Dec	14765	1476	10
23 Dec	10970	1568	7
27 Dec	10206	1458	7
28 Dec	9580	1597	6
29 Dec	10876	1554	7
30 Dec	9839	1406	7

"Trans" denotes the number of transactions, n is the sample size used to calculate the pre-averaged realized volatility, we have sampled every Sth transaction price, so the period over which returns are calculated is roughly 15 seconds.

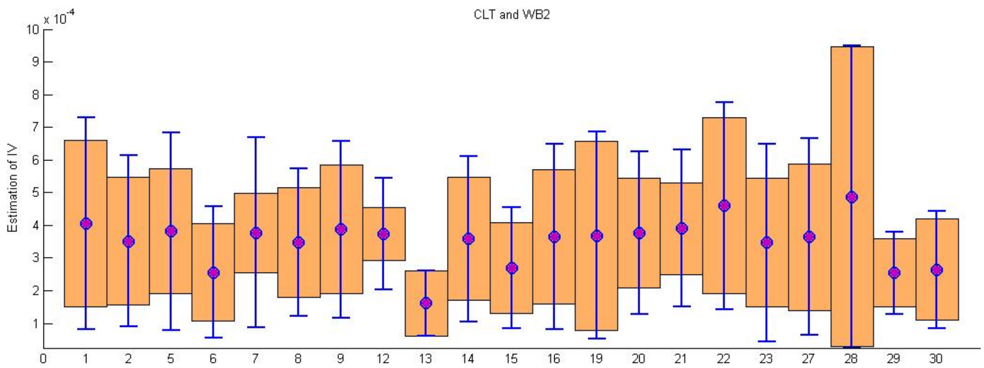

Figure 3.1. 95% Confidence Intervals (CI's) for the daily IV, for each regular exchange opening days in December 2011, calculated using the asymptotic theory of Podolskij and Vetter (2009) (CI's with bars), and the wild bootstrap method using WB2 as external random variable (CI's with lines). The pre-averaging realized volatility estimator is the middle of all CI's by construction. Days on the x-axis.

3.9.2 Appendix B

Proof of Theorem 1.3.1. For part (1), given that $\bar{Y}_j{}^* = \bar{Y}_j v_j$, where v_j are i.i.d. with $\mu_q^* = E^* |v_j|^q$, for any $q > 0$, we have that

$$\begin{aligned} V_n^* &= Var^* \left(n^{1/4} \frac{c_1 c_2}{\nu_1} \sum_{j=1}^{n/K} \bar{Y}_j{}^{*2} \right) = Var^* \left(n^{1/4} \frac{c_1 c_2}{\nu_1} \sum_{j=1}^{n/K} \bar{Y}_j^2 v_j^2 \right) \\ &= \left(\mu_4^* - (\mu_2^*)^2 \right) \frac{c_1^2 c_2^2}{\nu_1^2} n^{1/2} \sum_{j=1}^{n/K} \bar{Y}_j^4 = \hat{V}_n \end{aligned}$$

under the condition that $\mu_4^* - (\mu_2^*)^2 = \frac{2}{3}$. Thus,

$$V_n^* \xrightarrow{P} V = \frac{2 c_1^2 c_2^2}{\nu_1^2} \int_0^1 \left(\nu_1 \sigma_u^2 + \nu_2 \omega^2 \right)^2 du,$$

by an application of Theorem 1 of Podolskij and Vetter (2009) (where we set $r = 4$ and $l = 0$).

For part (2), let $S_n^* = \sum_{j=1}^{n/K} z_j^*$, where $z_j^* = \frac{c_1 c_2}{\nu_1} n^{1/4} \left(\bar{Y}_j{}^{*2} - E^* \left(\bar{Y}_j{}^{*2} \right) \right)$. Note that $E^* \left(z_j^* \right) = 0$ and that

$$Var^* \left(\sum_{j=1}^{n/K} z_j^* \right) = V_n^* \xrightarrow{P} V,$$

by part (1). Moreover, since $z_1^*, \ldots, z_{n/K}^*$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta > 0$ and for some constant $C > 0$,

$$\sup_{x \in \mathbb{R}} \left| P^* \left(S_n^* \leq x \right) - \Phi \left(x / \sqrt{V} \right) \right| \leq C \sum_{j=1}^{n/K} E^* \left| z_j^* \right|^{2+\delta},$$

which converges to zero in probability as $n \to \infty$. Indeed, we have that

$$\begin{aligned} \sum_{j=1}^{n/K} E^* \left| z_j^* \right|^{2+\delta} &= \left| \frac{c_1 c_2}{\nu_1} \right|^{2+\delta} \sum_{j=1}^{n/K} E^* \left| n^{1/4} \left(\bar{Y}_j^{*2} - E^* \left(\bar{Y}_j^{*2} \right) \right) \right|^{2+\delta} \\ &\leq 2 \left| \frac{c_1 c_2}{\nu_1} \right|^{2+\delta} n^{\frac{(2+\delta)}{4}} \sum_{j=1}^{n/K} E^* \left| \bar{Y}_j^{*2} \right|^{2+\delta} \\ &\leq 2 \left| \frac{c_1 c_2}{\nu_1} \right|^{2+\delta} E^* \left| v_1 \right|^{2(2+\delta)} n^{-\frac{\delta}{4}} \left(n^{\frac{(1+\delta)}{2}} \sum_{j=1}^{n/K} \left| \bar{Y}_j^2 \right|^{2+\delta} \right) \\ &= O_p \left(n^{-\frac{\delta}{4}} \right) = o_p(1), \end{aligned}$$

since $E\left|v_1\right|^{2(2+\delta)} \leq \Delta < \infty$ by assumption, and given that by Theorem 1 of Podolskij and Vetter (2009)

$$n^{\frac{(1+\delta)}{2}} \sum_{j=1}^{n/K} \left|\overline{Y}_j\right|^{2(2+\delta)} \xrightarrow{P} \frac{\mu_{2(2+\delta)}}{c_1 c_2} \int_0^1 \left(\nu_1 \sigma_u^2 + \nu_2 \omega^2\right)^{2+\delta} du,$$

which is bounded since σ is an adapted càdlàg spot volatility process and locally bounded away from zero, and $E(\epsilon_t^2) = \omega^2 < \Delta < \infty$.

Proof of Theorem 1.3.2 Given that $T_n \xrightarrow{d} N(0,1)$ (cf. Corollary 1 of Podolskij and Vetter (2009), it suffices to show that $T_n^* \xrightarrow{d^*} N(0,1)$ in probability. Let

$$H_n^* = \frac{n^{1/4}\left(PRV_n^* - E^*\left(PRV_n^*\right)\right)}{\sqrt{V_n^*}},$$

and note that

$$T_n^* = H_n^* \sqrt{\frac{V_n^*}{\hat{V}_n^*}},$$

where $\hat{V}_n^*$ is defined in the main text. Theorem 1.3.1 proved that $H_n^* \xrightarrow{d^*} N(0,1)$ in probability. Thus, it suffices to show that $\hat{V}_n^* - V_n^* \xrightarrow{P^*} 0$ in probability. In particular, we show that (1) $Bias^*\left(\hat{V}_n^*\right) = 0$, and (2) $Var^*\left(\hat{V}_n^*\right) \xrightarrow{P} 0$. It is easy to verify that (1) holds by the definition of $\hat{V}_n^*$ and V_n^*. To prove (2), note that

$$\begin{aligned}
Var^*\left(\hat{V}_n^*\right) &= E^*\left(\hat{V}_n^* - V_n^*\right)^2 = \left(\frac{\mu_4^* - \left(\mu_2^*\right)^2}{\mu_4^*}\right)^2 n \frac{c_1^2 c_2^2}{\nu_1^2} E^*\left(\sum_{j=1}^{n/K}\left(\overline{Y}_j^4 v_j^4 - \mu_4^* \overline{Y}_j^4\right)\right)^2 \\
&= \left(\frac{\mu_4^* - \left(\mu_2^*\right)^2}{\mu_4^*}\right)^2 n \frac{c_1^2 c_2^2}{\nu_1^2} \sum_{j=1}^{n/K} \overline{Y}_j^8 E^*\left(v_j^4 - \mu_4^*\right)^2 \\
&= \left(\frac{\mu_4^* - \left(\mu_2^*\right)^2}{\mu_4^*}\right)^2 \left(\mu_8^* - \mu_4^{*2}\right) \frac{c_1^2 c_2^2}{\nu_1^2} n^{-\frac{1}{2}} n^{\frac{3}{2}} \sum_{j=1}^{n/K} \overline{Y}_j^8 \\
&= O_P\left(n^{-\frac{1}{2}}\right) = o_P(1),
\end{aligned}$$

where we have used the independence of v_j over j to justify the third equality and Theorem 1 of Podolskij and Vetter (2009) (with $r = 8$ and $l = 0$) to justify the fact that $n^{\frac{3}{2}} \sum_{j=1}^{n/K} \overline{Y}_j^8 = O_P(1)$. This requires strengthening the moment condition on ϵ by assuming that $E\left|\epsilon\right|^{2(8+\varepsilon)} < \infty$.

3.10 Appendix for Chapter 2

3.10.1 Appendix C

Here we describe the Minimum Volatility Method algorithm of Politis, Romano and Wolf (1999, Chapter 9) for choosing the block size b_n for a two-sided confidence interval.

Algorithm: Choice of the bootstrap block size by minimizing confidence interval volatility

(i) For $b = b_{small}$ to $b = b_{big}$ compute a bootstrap interval for IV at the desired confidence level, this resulting in endpoints $IC_{b,low}$ and $IC_{b,up}$.

(ii) For each b compute the volatility index VI_b as the standard deviation of the interval endpoints in a neighborhood of b. More specifically, for a smaller integer d, let VI_b equal to the standard deviation of the endpoints $\{IC_{b-d,low}, \ldots, IC_{b+d,low}\}$ plus the standard deviation of the endpoints $\{IC_{b-d,up}, \ldots, IC_{b+d,up}\}$, i.e.

$$VI_b \equiv \sqrt{\frac{1}{2d+1}\sum_{i=-d}^{d}\left(IC_{b+i,low} - \bar{IC}_{low}\right)^2} + \sqrt{\frac{1}{2d+1}\sum_{i=-d}^{d}\left(IC_{b+i,up} - \bar{IC}_{up}\right)^2},$$

where $\bar{IC}_{low} = \frac{1}{2d+1}\sum_{i=-d}^{d} IC_{b+i,low}$ and $\bar{IC}_{up} = \frac{1}{2d+1}\sum_{i=-d}^{d} IC_{b+i,up}$.

(iii) Pick the value b^* corresponding to the smallest volatility index and report $\{IC_{b^*,low}, IC_{b^*,up}\}$ as the final confidence interval.

To make the algorithm more computationally efficient, we have skipped a number of b values in regular fashion between b_{small} and b_{big}. We have considered only the values of b such that $b = pk_n$ where p is a fixed integer. We employ $b_{small} = 2k_n$, $b_{big} = \min(\theta\frac{N_n}{4}, 12k_n)$ and $d = 2$.

Tables 3.4 and 3.5 report the actual coverage rates for the feasible asymptotic theory approach and for our bootstrap methods using the optimal block size by minimizing confidence interval volatility. In Table 3.6 we provide some statistics of GE shares in January 2011.

TABLE 3.4. Coverage rates of Nominal 95% intervals using $\theta = 1/3$

	n	SV1F CLT	SV1F Boot	SV1F Avg. Block size	SV2F CLT	SV2F Boot	SV2F Avg. Block size
$\xi^2 = 0.0001$							
	195	90.89	91.02	11.75	88.60	90.09	11.73
	390	91.52	91.74	21.09	90.32	90.98	22.16
	780	92.88	93.41	32.63	91.40	92.94	34.31
	1560	93.86	94.01	65.50	92.62	93.71	68.58
	4680	94.32	94.43	144.69	93.94	94.43	143.98
	7800	94.68	94.72	172.40	94.19	95.02	179.48
	11700	94.60	94.87	220.48	94.17	95.14	224.81
	23400	94.80	94.93	319.21	94.68	95.10	319.67
$\xi^2 = 0.001$							
	195	90.77	90.88	11.68	88.20	90.07	11.80
	390	91.14	91.43	20.71	89.31	90.21	21.78
	780	92.26	93.50	32.33	90.80	92.54	34.24
	1560	93.40	94.12	65.11	92.61	94.85	69.73
	4680	94.46	95.07	140.71	93.65	95.20	151.50
	7800	94.14	95.24	174.08	94.05	95.35	172.34
	11700	94.23	95.13	219.74	93.98	95.45	222.15
	23400	94.47	95.04	323.09	94.50	95.23	312.43
$\xi^2 = 0.01$							
	195	83.11	88.51	11.73	80.96	87.79	11.56
	390	84.45	91.16	20.68	83.91	89.98	21.81
	780	86.48	91.92	31.67	85.89	91.96	32.09
	1560	87.97	93.10	64.84	88.02	93.61	62.08
	4680	91.13	94.17	144.19	90.76	94.12	142.92
	7800	91.92	94.91	170.45	91.45	94.26	170.06
	11700	92.20	94.52	216.41	92.19	94.61	215.82
	23400	92.87	94.85	323.29	92.88	95.12	315.95

Notes: CLT-intervals based on the Normal; Boot-intervals based on the bootstrap. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 3.5. Coverage rates of Nominal 95% intervals using $\theta = 1$

		SV1F			SV2F		
	n	CLT	Boot	Avg. Block size	CLT	Boot	Avg. Block size
$\xi^2 = 0.0001$							
	195	89.48	90.10	35.90	84.50	87.12	36.27
	390	91.41	94.30	65.86	86.63	91.47	65.86
	780	92.81	94.98	132.22	88.71	92.10	124.99
	1560	93.57	95.13	262.24	90.39	93.92	235.76
	4680	94.19	95.45	517.12	91.50	94.20	451.02
	7800	94.27	95.12	682.52	92.76	95.00	594.68
	11700	94.06	95.50	804.62	93.15	94.81	713.17
	23400	94.39	95.48	1210.69	93.80	94.90	1063.81
$\xi^2 = 0.001$							
	195	89.05	92.19	35.90	84.41	87.60	35.92
	390	91.31	94.63	65.78	86.90	91.86	66.06
	780	92.96	94.76	132.78	88.57	92.80	124.15
	1560	93.66	95.37	265.00	90.34	94.30	237.96
	4680	94.12	95.52	514.43	92.03	94.51	458.33
	7800	94.21	95.16	688.04	92.32	94.88	582.40
	11700	94.17	95.18	806.15	92.98	95.01	719.93
	23400	94.35	95.11	1210.23	93.80	94.86	1062.43
$\xi^2 = 0.01$							
	195	88.42	92.18	35.81	84.07	88.62	35.97
	390	90.51	94.60	66.44	86.58	91.31	66.16
	780	92.17	95.12	132.58	88.52	92.87	125.22
	1560	93.35	95.15	264.96	90.01	94.40	243.92
	4680	93.77	95.60	515.74	91.72	95.23	471.10
	7800	94.28	95.72	671.84	92.76	95.20	593.08
	11700	94.16	95.24	808.00	93.03	95.40	732.35
	23400	94.26	95.18	1197.28	93.70	95.31	1081.40

Notes: CLT-intervals based on the Normal; Boot-intervals based on the bootstrap. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 3.6. Summary statistics

Days	Trans	S	n
3 Oct	12613	9	1402
4 Oct	13782	9	1532
5 Oct	10628	7	1519
6 Oct	9991	7	1428
7 Oct	9785	7	1398
10 Oct	10660	7	1523
11 Oct	8588	6	1432
12 Oct	11160	7	1595
13 Oct	8649	6	1442
14 Oct	9261	6	1544
17 Oct	8530	6	1422
18 Oct	8751	6	1459
19 Oct	9023	6	1504
20 Oct	9251	6	1542
21 Oct	12513	8	1565
24 Oct	11642	8	1456
25 Oct	10919	8	1365
26 Oct	9249	6	1542
27 Oct	14598	9	1622
28 Oct	9405	6	1568
31 Oct	8871	6	1500

"Trans" denotes the number of transactions, n is the sample size used to calculate the pre-averaged realized volatility, we have sampled every Sth transaction price, so the period over which returns are calculated is roughly 15 seconds.

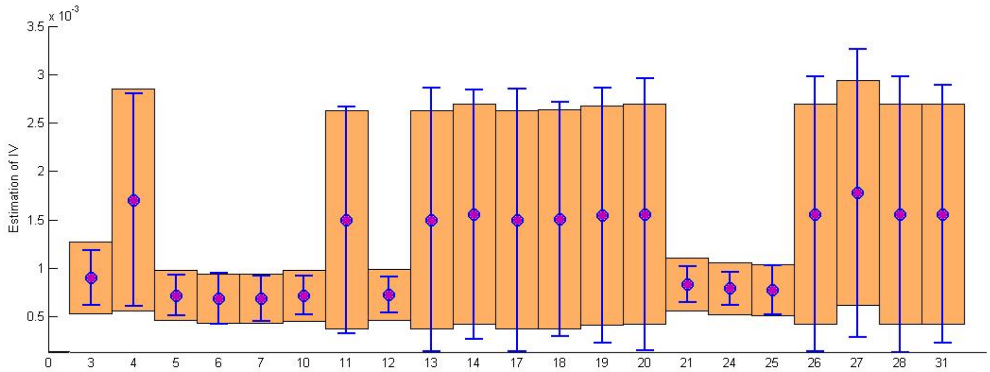

Figure 3.2. 95% Confidence Intervals (CI's) for the daily IV, for each regular exchange opening days in October 2011, calculated using the asymptotic theory of Jacod et al. (2009) (CI's with bars), and the wild blocks of blocks bootstrap method (CI's with lines). The pre-averaging realized volatility estimator is the middle of all CI's by construction. Days on the x-axis.

3.10.2 Appendix D: Proofs

As in Jacod et al. (2009), we assume throughout this Appendix that the processes a, σ and X are bounded processes satisfying (2.1) with a and σ adapted càdlàg processes. As Jacod et al. (2009) explain, this assumption simplifies the mathematical derivations without loss of generality (by a standard localization procedure detailed in Jacod (2008)). Formally, we derive our results under the following assumption.

Assumption 3. *X satisfies equation (2.1) with a and σ adapted càdlàg processes such that a, σ, and X are bounded processes (implying that α is also bounded).*

Notation

In the following, K denotes a constant which changes from line to line. Moreover, we follow Jacod et al. (2009) and use the following additional notation. We let

$$\bar{X}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(X_{\frac{i+j}{n}} - X_{\frac{i+j-1}{n}}\right), \qquad \bar{\epsilon}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(\epsilon_{\frac{i+j}{n}} - \epsilon_{\frac{i+j-1}{n}}\right),$$

and note that $\bar{Y}_i = \bar{X}_i + \bar{\epsilon}_i$. In addition, we let

$$\begin{aligned}
c_i &= \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)^2 \int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} \sigma_t^2 dt; \\
A_i &= E\left(\bar{\epsilon}_i^2\right) = \sum_{j=0}^{k_n-1}\left(g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right)\right)^2 E\left(\epsilon^2\right) = \frac{\psi_1^{k_n}}{k_n} E\left(\epsilon^2\right) \text{ and} \\
\widetilde{Y}_i &= \bar{Y}_i^2 - A_i - c_i.;
\end{aligned}$$

Following Jacod et al. (2009), we also introduce the following random variables. For $j = 1, \ldots, J_n$, we let

$$\eta\left(p\right)_j = \frac{1}{\theta\psi_2\sqrt{n}}\zeta(p)_{(j-1)(p+1)k_n}, \text{ with } \zeta\left(p\right)_j = \sum_{i=j}^{j+(p+1)k_n-1} \widetilde{Y}_i,$$

where $p \geq 1$ is a fixed integer; $\eta\left(p\right)_j$ is the normalized sum of squared pre-averaged returns $\widetilde{Y}_i$ over a block of size $b_n = (p+1)\,k_n$. Note that $\eta\left(p\right)_j$ is measurable with respect to $\mathcal{F}^n_{j(p+1)k_n}$, the sigma algebra generated by all $\mathcal{F}^0_{j(p+1)k_n/n}$-measurable random variables plus all variables Y_s, with $s < j\left(p+1\right)k_n$. Finally, we let

$$\beta(p)_i = \sup\nolimits_{s,t\in\left[\frac{i}{n}, \frac{i+(p+1)k_n}{n}\right]}\left(|a_s - a_t| + |\sigma_s - \sigma_t| + |\alpha_s - \alpha_t|\right), \tag{3.20}$$

and

$$\gamma^2(p)_t = \frac{4}{\psi_2^2}\left(\left(\Phi_{22} + \frac{1}{p+1}\Psi_{22}\right)\theta\sigma_t^4 + 2\left(\Phi_{12} + \frac{1}{p+1}\Psi_{12}\right)\frac{\sigma_t^2\alpha_t}{\theta} + \left(\Phi_{11} + \frac{1}{p+1}\Psi_{11}\right)\frac{\alpha_t^2}{\theta^3}\right). \tag{3.21}$$

Our bootstrap estimators depend crucially on

$$\bar{B}_j \equiv \frac{1}{b_n}\sum_{i=1}^{b_n} \bar{Y}_{i-1+(j-1)b_n}^2 = \frac{1}{b_n}\sum_{i=(j-1)b_n}^{jb_n-1} \bar{Y}_i^2, \text{ for } j = 1, \dots, J_n,$$

where $J_n = N_n/b_n$ is the number of non-overlapping blocks of size b_n out of $N_n = n-k_n+2$ observations on pre-averaged returns.

Our first result is instrumental in proving our bootstrap results.

Lemma 3.10.1. *Suppose Assumptions 2 and 3 hold. Then, for all integer $p \geq 1$, and each $q > 0$, we have that*

a1) $\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^q\right) \to 0.$

a2) $\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^q \to^P 0.$

a3) $\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^q | \mathcal{F}_{(j-1)(p+1)k_n}^n\right)\right) \to 0.$

a4) $\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^q | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) \to^P 0.$

a5) $\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}E\left(\beta\left(2p+1\right)_{(j-1)(p+1)k_n}^q | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) \to^P 0.$

a6) $\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right)} \to^P 0.$

a7) $\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(2p+1\right)_{(j-1)(p+1)k_n}^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right)} \to^P 0.$

Proof of Lemma 3.10.1. Part a1). Given the definition of $\beta\left(p\right)_{(j-1)(p+1)k_n}$ we can write

$$\begin{aligned}
\beta\left(p\right)_{(j-1)(p+1)k_n} &\leq \sup\nolimits_{s,t\in\left[\frac{(j-1)(p+1)k_n}{n},\frac{(j-1)(p+1)k_n+(p+1)k_n}{n}\right]}\left(|a_s - a_t|\right) \\
&\quad + \sup\nolimits_{s,t\in\left[\frac{(j-1)(p+1)k_n}{n},\frac{(j-1)(p+1)k_n+(p+1)k_n}{n}\right]}\left(|\sigma_s - \sigma_t|\right) \\
&\quad + \sup\nolimits_{s,t\in\left[\frac{(j-1)(p+1)k_n}{n},\frac{(j-1)(p+1)k_n+(p+1)k_n}{n}\right]}\left(|\alpha_s - \alpha_t|\right) \\
&\equiv \Gamma\left(a,p\right)_{(j-1)(p+1)k_n} + \Gamma\left(\sigma,p\right)_{(j-1)(p+1)k_n} + \Gamma\left(\alpha,p\right)_{(j-1)(p+1)k_n}.
\end{aligned}$$

Given that $\Gamma\left(a,p\right)_{(j-1)(p+1)k_n}, \Gamma\left(\sigma,p\right)_{(j-1)(p+1)k_n}$ and $\Gamma\left(\alpha,p\right)_{(j-1)(p+1)k_n}$ are strictly positive, for any $q > 0$, using the c-r inequality, we can write

$$\beta\left(p\right)_{(j-1)(p+1)k_n}^q \leq K\left(\Gamma\left(\sigma,p\right)_{(j-1)(p+1)k_n}^q + \Gamma\left(a,p\right)_{(j-1)(p+1)k_n}^q + \Gamma\left(\alpha,p\right)_{(j-1)(p+1)k_n}^q\right).$$

It follows that

$$n^{-1/2}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)^q_{(j-1)(p+1)k_n}\right)\leq Kn^{-1/2}E\left(\sum_{j=1}^{J_n}\Gamma\left(\sigma,p\right)^q_{(j-1)(p+1)k_n}\right)$$
$$+Kn^{-1/2}E\left(\sum_{j=1}^{J_n}\Gamma\left(a,p\right)^q_{(j-1)(p+1)k_n}\right)$$
$$+Kn^{-1/2}E\left(\sum_{j=1}^{J_n}\Gamma\left(\alpha,p\right)^q_{(j-1)(p+1)k_n}\right)=o\left(1\right),$$

where we use Lemma 5.3 of Jacod, Podolskij and Vetter (2010) to show that each of the terms above are $o\left(1\right)$ (given that a, σ and α are càdlàg bounded processes).

Proof of Lemma 3.10.1. Part a2). Note that given the result of part a1) of Lemma 3.10.1, it is sufficient to show that $\frac{1}{n}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)^q_{(j-1)(p+1)k_n}\right)^2\to 0$. By the c-r inequality,

$$\frac{1}{n}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)^q_{(j-1)(p+1)k_n}\right)^2\leq\frac{J_n}{n}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)^{2q}_{(j-1)(p+1)k_n}\right)\leq K\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)^{2q}_{(j-1)(p+1)k_n}\right),$$

which is $o\left(1\right)$ by part a1) of Lemma 3.10.1 and given that $J_n=O\left(\sqrt{n}\right)$.

Proof of Lemma 3.10.1. Part a3). Given the law of iterated expectations, the result follows directly from part a1) of Lemma **3.10.1.**

Proof of Lemma 3.10.1. Part a4). The proof follows similarly as in part a2) of Lemma **3.10.1**, where we now consider the variable $E\left(\beta\left(p\right)^q_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)$ in place of $\beta\left(p\right)^q_{(j-1)(p+1)k_n}$.

Proof of Lemma 3.10.1. Part a5). Given the definition of $\beta(p)_i$, for any $p\geq 1$, such that $b_n=\left(p+1\right)k_n$ we can write

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}E\left(\beta\left(2p+1\right)^q_{(j-1)b_n}|\mathcal{F}^n_{(j-1)b_n}\right) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{\left[\frac{J_n}{2}\right]}E\left(\beta\left(2p+1\right)^q_{2(j-1)b_n}|\mathcal{F}^n_{2(j-1)b_n}\right)\\ &+\frac{1}{\sqrt{n}}\sum_{j=1}^{\left[\frac{J_n}{2}\right]}E\left(\beta\left(2p+1\right)^q_{(2(j-1)+1)b_n}|\mathcal{F}^n_{(2(j-1)+1)b_n}\right),\end{aligned}$$

which is $o_P\left(1\right)$ given part a4) of Lemma 3.10.1.

Proof of Lemma 3.10.1. Part a6). Here, the proof contains two steps. Step 1. We show show that $\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)\to 0$. Step 2. We show show that $\frac{1}{n}Var\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)\to 0$. Note that using the first expression in equation (5.47) of Jacod et al. (2009), the result of step 1 follows directly. Given this result, to show step 2, it is sufficient to show that $\frac{1}{n}E\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)^2 \to 0$. We have that

$$\frac{1}{n}\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)^2 \leq \frac{J_n}{n}\sum_{j=1}^{J_n}E\left(E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right)$$
$$= \frac{J_n}{n}\sum_{j=1}^{J_n}E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^2\right) \leq K\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^2\right),$$

which is $o\left(1\right)$ given equation (5.47) of Jacod et al. (2009) and the fact that $J_n = O\left(\sqrt{n}\right)$ under our assumptions.

Proof of Lemma 3.10.1. Part a7). The proof follows similarly as part a5) and therefore we omit the details.

Our next result is crucial to the proofs of Lemmas 2.3.1 and 2.3.2.

Lemma 3.10.2. *Under Assumptions 1, 2 and 3, if* $b_n = (p+1)\,k_n$ *where* $p \geq 1$ *is fixed, then*

a1) $\frac{\sqrt{n}b_n^2}{k_n^2\psi_2^2}\sum_{j=1}^{J_n}\bar{B}_j^2 \to^P V_p + \theta\left(p+1\right)\int_0^1\left(\sigma_s^2 + \frac{\psi_1}{\theta^2\psi_2}\alpha_s\right)^2 ds.$

a2) $\frac{\sqrt{n}b_n^2}{k_n^2\psi_2^2}\sum_{j=1}^{J_n-1}\bar{B}_j\bar{B}_{j+1} \to^P \theta\left(p+1\right)\int_0^1\left(\sigma_s^2 + \frac{\psi_1}{\theta^2\psi_2}\alpha_s\right)^2 ds + O_P\left(\frac{1}{p}\right).$

Proof of Lemma 3.10.2. Part a1). Given the definition of $\bar{B}_j$, we have that

$$\bar{B}_j = \frac{1}{b_n}\sum_{i=(j-1)b_n}^{jb_n-1}\bar{Y}_i^2 = \frac{1}{b_n}\sum_{i=(j-1)b_n}^{jb_n-1}\underbrace{\left(\bar{Y}_i^2 - A_i - c_i\right)}_{\equiv\tilde{Y}_i} + \frac{1}{b_n}\sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i + c_i\right)$$

where $A_i \equiv E\left(\bar{\epsilon}_i^2\right)$ and $c_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)^2\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\sigma_t^2 dt$. It follows that

$$\frac{\sqrt{n}b_n^2}{k_n^2\psi_2^2}\sum_{j=1}^{J_n}\bar{B}_j^2 = \mathcal{B}_{1n} + \mathcal{B}_{2n} + \mathcal{B}_{3n},$$

where

$$\begin{aligned}
\mathcal{B}_{1n} &\equiv \sqrt{n}\sum_{j=1}^{J_n}\left(\frac{1}{\theta\psi_2\sqrt{n}}\sum_{i=(j-1)b_n}^{jb_n-1}\widetilde{Y}_i\right)^2 = \sqrt{n}\sum_{j=1}^{J_n}\eta\left(p\right)_j^2, \\
\mathcal{B}_{2n} &\equiv \frac{2}{\theta\psi_2}\sum_{j=1}^{J_n}\eta\left(p\right)_j\sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i+c_i\right);\ \text{and} \\
\mathcal{B}_{3n} &\equiv \frac{1}{\theta^2\psi_2^2\sqrt{n}}\sum_{j=1}^{J_n}\left(\sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i+c_i\right)\right)^2.
\end{aligned}$$

We show that (1) $\mathcal{B}_{1n}\to^P\int_0^1\gamma_t^2\left(p\right)dt$; (2) $\mathcal{B}_{2n}\to^P 0$, and that (3)

$$\mathcal{B}_{3n}\to^P\left(p+1\right)\theta\int_0^1\left(\sigma_t^2+\frac{\psi_1}{\theta^2\psi_2}\alpha_t\right)^2dt.$$

Starting with (1), write

$$\begin{aligned}
\sqrt{n}\sum_{j=1}^{J_n}\eta\left(p\right)_j^2-\int_0^1\gamma_t^2\left(p\right)dt &= \mathcal{B}_{1.1n}+\mathcal{B}_{1.2n}+\mathcal{B}_{1.3n}, \quad \text{with} \\
\mathcal{B}_{1.1n} &= \sqrt{n}\sum_{j=1}^{J_n}\left(\eta\left(p\right)_j^2-E\left(\eta\left(p\right)_j^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right), \\
\mathcal{B}_{1.2n} &= \sqrt{n}\sum_{j=1}^{J_n}E\left(\eta\left(p\right)_j^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)-\frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\gamma(p)^2_{\frac{j-1}{J_n}}, \\
\mathcal{B}_{1.3n} &= \frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\gamma(p)^2_{\frac{j-1}{J_n}}-\int_0^1\gamma_t^2\left(p\right)dt.
\end{aligned}$$

We show that each of $\mathcal{B}_{1.\ell n}\to^P 0$ for $\ell=1,2,3$. Starting with $\ell=1$, by Lenglart's inequality (see e.g. Lemma 4.4 of Vetter 2008), it is sufficient to show that $n\sum_{j=1}^{J_n}E\left(\eta\left(p\right)_j^4|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\to^P 0$, which follows immediately by using equation (5.57) of Jacod et al. (2009). This shows that $\mathcal{B}_{1.1n} \to^P 0$. Next, to show that $\mathcal{B}_{1.2n} \to^P 0$, note that

$$\begin{aligned}
\mathcal{B}_{1.2n} &\leq \sum_{j=1}^{J_n} \left| \sqrt{n} E\left(\eta\left(p\right)_j^2 | \mathcal{F}^n_{(j-1)(p+1)k_n}\right) - \frac{N_n}{n}\frac{1}{J_n}\gamma(p)^2_{\frac{j-1}{J_n}} \right| \\
&= \sum_{j=1}^{J_n} \left| \sqrt{n} E\left(\frac{1}{\theta^2\psi_2^2 n}\zeta^2(p)_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n}\right) - \frac{1}{n}\left(p+1\right)\theta\sqrt{n}\gamma(p)^2_{\frac{j-1}{J_n}} \right| \\
&= \frac{\sqrt{n}}{\theta^2\psi_2^2 n}\sum_{j=1}^{J_n} \left| E\left(\zeta^2(p)_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n}\right) - \theta^3\psi_2^2\left(p+1\right)\gamma(p)^2_{\frac{j-1}{J_n}} \right| \\
&\leq \frac{K_p}{\theta^2\psi_2^2\sqrt{n}}\sum_{j=1}^{J_n}\chi(p)_{(j-1)(p+1)k_n}
\end{aligned}$$

where the first line follows by the triangle inequality; the second line uses the definitions $\eta\left(p\right)_j = \frac{1}{\theta\psi_2\sqrt{n}}\zeta(p)_{(j-1)(p+1)k_n}$ and $N_n/J_n = \left(p+1\right)k_n$ with $k_n = \theta\sqrt{n}$; and the fourth line uses equation (5.41) of Jacod et al. (2009) to bound the term in absolute value, where

$$\chi(p)_{(j-1)(p+1)k_n} = n^{-1/4} + \sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n}\right)}$$

and $\beta\left(p\right)_i$ is as defined in (3.20). It follows that,

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\chi(p)_{(j-1)(p+1)k_n} \leq \frac{1}{\sqrt{n}}\sum_{j=1}^{J_n} n^{-1/4} + \frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n}\right)} \to^P 0,$$

where the first term is of order $O\left(n^{-1/4}\right)$ and the second term is $o_P\left(1\right)$ given part a6) of Lemma 3.10.1. Finally, $\mathcal{B}_{1.3n} \to^P 0$ follows immediately by Riemann's integrability of σ, the fact that $\frac{N_n}{n} \to 1$ and $J_n \to \infty$ as $n \to \infty$.

To show (2), let $\varphi_j \equiv \sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i + c_i\right)$ and $\zeta\left(X,p\right)_j = \sum_{i=(j-1)b_n}^{jb_n-1}\left(\bar{X}_i^2 - c_i\right)$. We can write

$$\begin{aligned}
\mathcal{B}_{2n} &= \frac{2}{\theta\psi_2}\sum_{j=1}^{J_n}\varphi_j \cdot \eta\left(p\right)_j = \mathcal{B}_{2.1n} + \mathcal{B}_{2.2n}, \quad \text{with} \\
\mathcal{B}_{2.1n} &= \frac{2}{\theta\psi_2}\sum_{j=1}^{J_n}\left(\varphi_j\eta\left(p\right)_j - E\left(\varphi_j\eta\left(p\right)_j | \mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right), \text{ and} \\
\mathcal{B}_{2.2n} &= \frac{2}{\theta\psi_2}\sum_{j=1}^{J_n}E\left(\varphi_j\eta\left(p\right)_j | \mathcal{F}^n_{(j-1)(p+1)k_n}\right).
\end{aligned}$$

We show that each of $\mathcal{B}_{2.\ell n} \to^P 0$ for $\ell = 1, 2$. Note that given the definitions of A_i, c_i, and the fact that $k_n = \theta\sqrt{n}$, Assumption 3 implies that $A_i + c_i \leq K/\sqrt{n}$ uniformly in i. Given that $b_n = \left(p+1\right)k_n$, it follows that $\varphi_j \leq K$ uniformly in j. Starting with $\ell = 1$, by Lenglart's inequality, it is sufficient to show that $\sum_{j=1}^{J_n} E\left(\varphi_j^2 \eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) \to^P 0$. We can write

$$\begin{aligned}
\sum_{j=1}^{J_n} E\left(\varphi_j^2 \eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) &\leq K \sum_{j=1}^{J_n} E\left(\eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) \\
&= K\left(\frac{1}{\sqrt{n}}\left(\sqrt{n}\sum_{j=1}^{J_n} E\left(\eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) - \frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\gamma(p)^2_{\frac{j-1}{J_n}}\right)\right) \\
&\quad + K\left(\frac{1}{\sqrt{n}}\left(\frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\gamma(p)^2_{\frac{j-1}{J_n}} - \int_0^1 \gamma_t^2\left(p\right)dt\right) + \frac{1}{\sqrt{n}}\int_0^1 \gamma_t^2\left(p\right)dt\right) \\
&\equiv K\left(\frac{1}{\sqrt{n}}\mathcal{B}_{1.2n} + \frac{1}{\sqrt{n}}\mathcal{B}_{1.3n} + \frac{1}{\sqrt{n}}\int_0^1 \gamma_t^2\left(p\right)dt\right) \\
&= \frac{1}{\sqrt{n}}o_P\left(1\right) + \frac{1}{\sqrt{n}}o_P\left(1\right) + O_P\left(\frac{1}{\sqrt{n}}\right) = o_P\left(1\right),
\end{aligned}$$

where in particular we use the fact that $\mathcal{B}_{1.2n} = o_P\left(1\right)$ and $\mathcal{B}_{1.3n} = o_P\left(1\right)$, and $\int_0^1 \gamma_t^2\left(p\right)dt = O_P\left(1\right)$. It follows that $\mathcal{B}_{2.1n} \to^P 0$. Next, to show that $\mathcal{B}_{2.2n} \to^P 0$, note that we can write

$$\mathcal{B}_{2.2n} \leq \frac{2K}{\theta\psi_2}\frac{1}{n^{1/4}}\left(n^{1/4}\sum_{j=1}^{J_n} E\left(\eta\left(p\right)_j | \mathcal{F}_{(j-1)(p+1)k_n}^n\right)\right) = O_P\left(n^{-1/4}\right)o_P\left(1\right) = o_P\left(1\right),$$

given that $\varphi_j \leq K$, and given equation (5.49) of Jacod et al. (2009).

Finally, to show (3), note that given the definitions of A_i and c_i, and by using equations (5.23) and (5.36) of Jacod et al. (2009), we can write

$$\sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i + c_i\right) = \sum_{i=(j-1)b_n}^{jb_n-1}\left(\frac{\psi_1}{\theta\sqrt{n}}\alpha_{(j-1)b_n/n} + \frac{\theta\psi_2}{\sqrt{n}}\sigma^2_{(j-1)b_n/n}\right) + O\left(\frac{p}{\sqrt{n}} + p\beta(p)_{(j-1)b_n}\right). \tag{3.22}$$

It follows that

$$\mathcal{B}_{3n} \equiv \frac{1}{\theta^2\psi_2^2\sqrt{n}}\sum_{j=1}^{J_n}\left(\sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i + c_i\right)\right)^2 = L_n + R_n,$$

where the leading term is

$$L_n = \left(p+1\right)\theta\frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\left(\frac{\psi_1}{\theta^2\psi_2}\alpha_{(j-1)b_n/n} + \sigma^2_{(j-1)b_n/n}\right)^2 \to^P (p+1)\theta\int_0^1\left(\sigma_t^2 + \frac{\psi_1}{\theta^2\psi_2}\alpha_t\right)^2 dt. \tag{3.23}$$

The remainder is such that

$$R_n = K \cdot O_P \left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} \beta(p)^2_{(j-1)b_n} \right) \to^P 0$$

by using Lemma (5.4) of Jacod et al. (2009).

Proof of part a2). Recall that $\bar{B}_j = \frac{1}{b_n} \sum_{i=(j-1)b_n}^{jb_n-1} \bar{Y}_i^2$ is the average of observations in a block of size b_n starting at observation $(j-1)\,b_n$. For any integer a_n such that $2 \leq a_n < b_n$, we can decompose

$$\bar{B}_j = \bar{B}_j^{[0,a_n-1]} + \bar{B}_j^{[a_n,b_n-1]},$$

where $\bar{B}_j^{[0,a_n-1]} \equiv \frac{1}{b_n} \sum\limits_{i=(j-1)b_n}^{(j-1)b_n+a_n-1} \bar{Y}_i^2$ and $\bar{B}_j^{[a_n,b_n-1]} \equiv \frac{1}{b_n} \sum\limits_{i=(j-1)b_n+a_n}^{jb_n-1} \bar{Y}_i^2$. Then

$$\begin{aligned}
\bar{B}_j \bar{B}_{j+1} &= \left(\bar{B}_j^{[0,pk_n-1]} + \bar{B}_j^{[pk_n,b_n-1]}\right)\left(\bar{B}_{j+1}^{[0,k_n-1]} + \bar{B}_{j+1}^{[k_n,b_n-1]}\right) \\
&= \left(\bar{B}_j^{[0,pk_n-1]} \bar{B}_{j+1}^{[0,k_n-1]}\right) + \left(\bar{B}_j^{[0,pk_n-1]} \bar{B}_{j+1}^{[k_n,b_n-1]}\right) + \left(\bar{B}_j^{[pk_n,b_n-1]} \bar{B}_{j+1}^{[k_n,b_n-1]}\right) \\
&+ \left(\bar{B}_j^{[pk_n,b_n-1]} \bar{B}_{j+1}^{[0,k_n-1]}\right) \\
&\equiv \underbrace{\Xi_{1j} + \Xi_{2j} + \Xi_{3j}}_{\equiv L_j} + \Xi_{4j}. \qquad (3.24)
\end{aligned}$$

We can write

$$\frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \bar{B}_j \bar{B}_{j+1} = \frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} L_j + \frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}.$$

The proof contains two steps. Step 1. We show that $\frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum\limits_{j=1}^{J_n-1} L_j \to^P (p+1)\theta \int_0^1 \left(\sigma_t^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_t\right)^2 dt$. Step 2. We show that $\frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum\limits_{j=1}^{J_n-1} \Xi_{4j} = O_P\left(\frac{k_n}{b_n}\right)$.

Step 1. Let $\overline{\sigma\epsilon}_j \equiv \frac{\psi_2 k_n}{n} \sigma^2_{(j-1)b_n/n} + \frac{\psi_1}{k_n} \alpha_{(j-1)b_n/n}$. It follows that

$$\frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} L_j - \left((p+1) - \frac{1}{p+1}\right) \theta \int_0^1 \left(\sigma_t^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_t\right)^2 dt = \mathcal{B}_{a.1n} + \mathcal{B}_{a.2n} + \mathcal{B}_{a.3n}, \text{ with}$$

$$\mathcal{B}_{a.1n} = \frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\sum_{j=1}^{J_n-1}\left(L_j - E\left(L_j|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right),$$
$$\mathcal{B}_{a.2n} = \frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\sum_{j=1}^{J_n-1}E\left(L_j|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) - \frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\sum_{j=1}^{J_n}\overline{\sigma\epsilon}_j^2,$$
$$\mathcal{B}_{a.3n} = \frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\sum_{j=1}^{J_n}\overline{\sigma\epsilon}_j^2 - \left((p+1) - \frac{1}{p+1}\right)\theta\int_0^1\left(\sigma_t^2 + \frac{\psi_1}{\theta^2\psi_2}\alpha_t\right)^2 dt.$$

We show that each of $\mathcal{B}_{a.\ell n} \to^P 0$ for $\ell = 1,2,3$. Starting with $\ell = 1$, by Lenglart's inequality, it is sufficient to show that $\frac{nb_n^4}{\psi_2^4 k_n^4}\sum_{j=1}^{J_n-1}E\left(L_j^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \to^P 0$. We can write

$$\begin{aligned}
\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\sum_{j=1}^{J_n-1}E\left(L_j^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) &= \frac{nb_n^4}{\psi_2^4 k_n^4}\sum_{j=1}^{J_n-1}E\left(\left(\bar{B}_j\bar{B}_{j+1} - \Xi_{4j}\right)^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \\
&\leq \frac{nb_n^4}{\psi_2^4 k_n^4}\sum_{j=1}^{J_n-1}E\left(\bar{B}_j^2\bar{B}_{j+1}^2|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \\
&\leq \frac{nb_n^4}{\psi_2^4 k_n^4}\sum_{j=1}^{J_n-1}\left(E\left(\bar{B}_j^4|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right)^{1/2}\left(E\left(\bar{B}_{j+1}^4|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right)^{1/2} \\
&= O_P\left(n^{-1/2}\right) = o_P(1),
\end{aligned}$$

where the first line uses the definition of L_j; the second line follows by the fact that $\Xi_{4j} \geq 0$; the third line follows by Cauchy-Schwartz inequality and the fourth line uses the fact that $b_n = (p+1)k_n$, $J_n = O(\sqrt{n})$ and $E\left(\bar{B}_j^4|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) = O_P(n^{-2})$ uniformly in j. To show that $E\left(\bar{B}_j^4|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) = O_P(n^{-2})$, by the c-r inequality,

$$\begin{aligned}
E\left(\bar{B}_j^4|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) &= \frac{1}{b_n^4}E\left(\left(\sum_{i=(j-1)b_n}^{jb_n}\bar{Y}_i^2\right)^4|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \\
&\leq \frac{1}{b_n}E\left(\left(\sum_{i=(j-1)b_n}^{jb_n}\bar{Y}_i^8\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \\
&\leq Kn^{-2},
\end{aligned}$$

where we can show that

$$E\left(\bar{Y}_i^8\right) \leq K\left(E\left(\bar{X}_i^8\right) + E\left(\bar{\epsilon}_i^8\right)\right) = O\left(n^{-2}\right) \text{ uniformly in i.}$$

given that $\bar{Y}_i = \bar{X}_i + \bar{\epsilon}_i$ and given equations (5.28) and (5.38) of Jacod et al. (2009). This shows that $\mathcal{B}_{a.1n} \to^P 0$. Next, to show that $\mathcal{B}_{a.2n} \to^P 0$, note that given the definition of

L_j, the fact that $b_n = (p+1)\,k_n$, and by using equation (3.24)we can write

$$\begin{aligned}
\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\sum_{j=1}^{J_n-1} E\left(L_j|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) &= n^{1/2}\frac{(p+1)^2}{\psi_2^2}\sum_{j=1}^{J_n-1} E\left(\left(\bar{B}_j^{[0,pk_n-1]}\bar{B}_{j+1}^{[0,k_n-1]}\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\\
&+n^{1/2}\frac{(p+1)^2}{\psi_2^2}\sum_{j=1}^{J_n-1} E\left(\left(\bar{B}_j^{[0,pk_n-1]}\bar{B}_{j+1}^{[k_n,b_n-1]}\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\\
&+n^{1/2}\frac{(p+1)^2}{\psi_2^2}\sum_{j=1}^{J_n-1} E\left(\left(\bar{B}_j^{[pk_n,b_n-1]}\bar{B}_{j+1}^{[k_n,b_n-1]}\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\\
&\equiv \Upsilon_1+\Upsilon_2+\Upsilon_3.
\end{aligned}$$

For Υ_1, we obtain

$$\begin{aligned}
\Upsilon_1 &= n^{1/2}\frac{(p+1)^2}{\psi_2^2}\sum_{j=1}^{J_n-1} E\left(\bar{B}_j^{[0,pk_n-1]}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)E\left(\bar{B}_{j+1}^{[0,k_n-1]}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\\
&= n^{1/2}\frac{(p+1)^2}{\psi_2^2 b_n^2}\sum_{j=1}^{J_n-1} E\left(\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}\bar{Y}_i^2\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)E\left(\left(\sum_{i=jb_n}^{jb_n+k_n-1}\bar{Y}_i^2\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right).
\end{aligned}$$

where we used the fact that $\bar{Y}_i$ and $\bar{Y}_j$ are (conditionally) independent provided that $|j-i|>k_n$. By adding and substracting appropriately, we can write $\bar{Y}_i^2=\left(\bar{Y}_i^2-c_i-A_i\right)+(c_i+A_i)$. Then we show that

$$E\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}\left(\bar{Y}_i^2-c_i-A_i\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) = pk_n\varpi_j \text{ and} \tag{3.25}$$

$$E\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}(c_i+A_i)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) = pk_n(\overline{\sigma\epsilon}_j+\varpi_j), \tag{3.26}$$

where $\varpi_j \equiv O_P\left(\frac{1}{n}+\frac{1}{\sqrt{n}}\sqrt{E\left(\beta\left(2p+1\right)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)$ and $\overline{\sigma\epsilon}_j \equiv \frac{\psi_2 k_n}{n}\sigma^2_{(j-1)b_n/n}+\frac{\psi_1}{k_n}\alpha_{(j-1)b_n/n}$. Given the decomposition $\bar{Y}_i^2=\left(\bar{X}_i+\bar{\epsilon}_i\right)^2=\bar{X}_i^2+2\bar{X}_i\bar{\epsilon}_i+\bar{\epsilon}_i^2$, we can write

$$\begin{aligned}
E\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}\left(\bar{Y}_i^2-c_i-A_i\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) &= E\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}\left(\bar{X}_i^2-c_i\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\\
&+ E\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}\left(2\bar{X}_i\bar{\epsilon}_i+\bar{\epsilon}_i^2-A_i\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\\
&\equiv \zeta_1(X)+\zeta_1(X,\epsilon)\,.
\end{aligned}$$

We can show that $\zeta_1(X,\epsilon)=0$ by relying on Assumption 1. In particular, noting that $\mathcal{F}^n_{(j-1)(p+1)k_n}\subset\mathcal{F}^0\times\mathcal{F}^1_{\frac{(j-1)(p+1)k_n}{n}-}$, (where $\mathcal{F}^0\times\mathcal{F}^1_{\frac{(j-1)(p+1)k_n}{n}-}$ denotes the sigma algebra generated by all $\mathcal{F}^0$-measurable random variables plus all variables Y_s, with $s < (j-1)(p+1)k_n)$ we have that by the law of iterated expectations,

$$\zeta_1(X,\epsilon)=E\left(E\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}\left(2\bar{X}_i\bar{\epsilon}_i+\bar{\epsilon}_i^2-A_i\right)|\mathcal{F}^0\times\mathcal{F}^1_{\frac{(j-1)(p+1)k_n}{n}-}\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)=0,$$

where by Assumption 1, $E\left(\bar{X}_i\bar{\epsilon}_i|\mathcal{F}^0\times\mathcal{F}^1_{\frac{(j-1)(p+1)k_n}{n}-}\right)=\bar{X}_iE\left(\bar{\epsilon}_i|\mathcal{F}^0\times\mathcal{F}^1_{\frac{(j-1)(p+1)k_n}{n}-}\right)=\bar{X}_iE\left(\bar{\epsilon}_i|X\right)=0$ and $E\left(\bar{\epsilon}_i^2|\mathcal{F}^0\times\mathcal{F}^1_{\frac{(j-1)(p+1)k_n}{n}-}\right)=E\left(\bar{\epsilon}_i^2|X\right)\equiv A_i$ (see equation (5.37) of Jacod et al. (2009)). For $\zeta_1(X)$, by the definition of c_i, we can write

$$\begin{aligned}
\zeta_1(X) &\le \frac{K}{n^{1/4}}\chi(p)_{(j-1)(p+1)k_n}\\
&= K\left(\frac{1}{n^{1/2}}+\frac{1}{n^{1/4}}\sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)\\
&\le \left(\frac{1}{n^{1/2}}+\frac{1}{n^{1/4}}\sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)\\
&= pk_n\varpi_j,
\end{aligned}$$

where the first line follows from equation (5.30) of Jacod et al. (2009); the second line uses the definition of $\chi(p)_{(j-1)(p+1)k_n}=n^{-1/4}+\sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}$; and the third line uses the fact that $\beta\left(p\right)_{(j-1)(p+1)k_n}\le\beta\left(2p+1\right)_{(j-1)(p+1)k_n}$. This proves (3.25). To show (3.26), we rely on arguments similar to those used by Jacod et al. (2009) (in particular, see their equations (5.23) and (5.36)). This implies that

$$E\left(\left(\sum_{i=(j-1)b_n}^{(j-1)b_n+pk_n-1}\bar{Y}_i^2\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)=pk_n\left(\overline{\sigma\epsilon}_j+2\varpi_j\right).$$

By similar arguments, we can show that

$$E\left(\left(\sum_{i=jb_n}^{jb_n+k_n-1}\bar{Y}_i^2\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)=k_n\left(\overline{\sigma\epsilon}_j+2\varpi_j\right),$$

which implies that

$$\begin{aligned}
\Upsilon_1 &= n^{1/2}\frac{(p+1)^2}{\psi_2^2b_n^2}\sum_{j=1}^{J_n-1}pk_n^2\left(\overline{\sigma\epsilon}_j+2\varpi_j\right)^2=n^{1/2}\frac{(p+1)^2}{\psi_2^2b_n^2}pk_n^2\sum_{j=1}^{J_n-1}\left(\overline{\sigma\epsilon}_j^2+4\varpi_j\overline{\sigma\epsilon}_j+2\varpi_j^2\right)\\
&= n^{1/2}\frac{p}{\psi_2^2}\sum_{j=1}^{J_n-1}\left(\overline{\sigma\epsilon}_j^2+4\varpi_j\overline{\sigma\epsilon}_j+2\varpi_j^2\right).
\end{aligned} \tag{3.27}$$

Using the fact that $\sqrt{n}\overline{\sigma\epsilon}_j = O_P(1)$ uniformly in j, $J_n = O(\sqrt{n})$, and the definition of $\varpi_j = O_P\left(\frac{1}{n} + \frac{1}{\sqrt{n}}E\left(\left(\beta(2p+1)_{(j-1)(p+1)k_n}\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right)$ uniformly in j, we get that

$$\begin{aligned}
n^{1/2}\frac{2p}{\psi_2^2}\sum_{j=1}^{J_n-1}\varpi_j\overline{\sigma\epsilon}_j &= \frac{2p}{\psi_2^2}\sum_{j=1}^{J_n-1}\varpi_j\left(n^{1/2}\overline{\sigma\epsilon}_j\right)\\
&= O_P\left(\frac{J_n}{n}\right) + O_P\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\sqrt{E\left(\beta(2p+1)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right),
\end{aligned}$$

which is $o_P(1)$ since $J_n/n = O(1/\sqrt{n})$ and $\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n-1}E\left(\left(\beta(2p+1)_{(j-1)(p+1)k_n}\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \to^P 0$ by Lemma 3.10.1. The third term in (3.27) is such that

$$\begin{aligned}
n^{1/2}\frac{2p}{\psi_2^2}\sum_{j=1}^{J_n-1}\varpi_j^2 &= O_P\left(\frac{J_n}{n^{3/2}}\right) + O_P\left(\frac{2}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\sqrt{E\left(\beta(2p+1)^2_{(j-1)(p+1)k_n}|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)}\right)\\
&\quad + O_P\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n-1}E\left(\left(\beta(2p+1)^2_{(j-1)(p+1)k_n}\right)|\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right) = o_P(1),
\end{aligned}$$

given parts a5) and a7) of Lemma 3.10.1. Thus

$$\Upsilon_1 = n^{1/2}\frac{p}{\psi_2^2}\sum_{j=1}^{J_n-1}\overline{\sigma\epsilon}_j^2 + o_P(1). \tag{3.28}$$

Similarly, we can show

$$\Upsilon_2 = n^{1/2}\frac{p^2}{\psi_2^2}\sum_{j=1}^{J_n-1}\overline{\sigma\epsilon}_j^2 + o_P(1), \text{ and} \tag{3.29}$$

$$\Upsilon_3 = n^{1/2}\frac{1}{\psi_2^2}\sum_{j=1}^{J_n-1}\overline{\sigma\epsilon}_j^2 + o_P(1). \tag{3.30}$$

From (3.28), (3.29) and (3.30), we have that

$$\begin{aligned}
\frac{n^{1/2}b_n^2}{\psi_2^2k_n^2}\sum_{j=1}^{J_n-1}E\left(L_j|\mathcal{F}^n_{(j-1)(p+1)k_n}\right) &= \left((p+1)^2-1\right)\frac{n^{1/2}}{\psi_2^2}\sum_{j=1}^{J_n-1}\overline{\sigma\epsilon}_j^2 + o_P(1)\\
&= \left((p+1)^2-1\right)\left(\frac{n^{1/2}}{\psi_2^2}\sum_{j=1}^{J_n}\overline{\sigma\epsilon}_j^2 - \frac{n^{1/2}}{\psi_2^2}\overline{\sigma\epsilon}_{J_n}^2\right) + o_P(1)\\
&= \left((p+1)^2-1\right)\left(\frac{n^{1/2}}{\psi_2^2}\sum_{j=1}^{J_n}\overline{\sigma\epsilon}_j^2\right) + O_P\left(n^{-1/2}\right) + o_P(1)\\
&= \left((p+1)^2-1\right)\left(\frac{n^{1/2}}{\psi_2^2}\sum_{j=1}^{J_n}\overline{\sigma\epsilon}_j^2\right) + o_P(1).
\end{aligned}$$

This shows that $\mathcal{B}_{a.2n} \to^P 0$. Finally, $\mathcal{B}_{a.3n} \to^P 0$ follows immediately by Riemann's integrability of α and σ, the fact that $\frac{N_n}{n} \to 1$ and $J_n \to \infty$ as $n \to \infty$.

Step 2. Next, we analyze the term that depends on $\Xi_{4j} \equiv \bar{B}_j^{[pk_n,b_n-1]} \bar{B}_{j+1}^{[0,k_n-1]}$. We show that $E\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}\right) = O\left(\frac{k_n}{b_n}\right)$, and $Var\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}\right) = O\left(\frac{k_n^2}{b_n^2}\right)$. Given the definition of Ξ_{4j}, we have that

$$
\begin{aligned}
E\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}\right) &= \frac{n^{1/2}}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} E\left(\left(\sum_{i=1}^{k_n} \bar{Y}^2_{i-1+(j-1)b_n+pk_n}\right)\left(\sum_{l=1}^{k_n} \bar{Y}^2_{l-1+jb_n}\right)\right) \\
&= \frac{n^{1/2}}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \sum_{i=1}^{k_n} \sum_{l=1}^{k_n} E\left(\bar{Y}^2_{i-1+(j-1)b_n+pk_n} \bar{Y}^2_{l-1+jb_n}\right) \\
&\leq \frac{n^{1/2}}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \sum_{i=1}^{k_n} \sum_{l=1}^{k_n} \left(E\left(\bar{Y}^4_{i-1+(j-1)b_n+pk_n}\right)\right)^{1/2} \left(E\left(\bar{Y}^4_{l-1+jb_n}\right)\right)^{1/2} \quad (3.31)
\end{aligned}
$$

Given the decomposition $\bar{Y}_i = \bar{X}_i + \bar{\epsilon}_i$, using the triangle inequality we have that $\left|\bar{Y}_i\right| \leq \left|\bar{X}_i\right| + |\bar{\epsilon}_i|$. It follows that

$$
E\left(\bar{Y}_i^4\right) \leq K\left(E\left(\bar{X}_i^4\right) + E\left(\bar{\epsilon}_i^4\right)\right) = O\left(n^{-1}\right)
$$

uniformly in i, where we use the c-r inequality and equations (5.28) and (5.38) of Jacod et al. (2009) to show that $E\left(\bar{X}_i^4\right) = O\left(n^{-1}\right)$, and $E\left(\bar{\epsilon}_i^4\right) = O\left(n^{-1}\right)$ uniformly in i. Thus, we can bound (3.31) by $K\frac{n^{1/2}}{\psi_2^2 k_n^2}\frac{J_n k_n^2}{n} = O\left(\frac{k_n}{b_n}\right)$ given that $J_n = N_n/b_n = \left(\sqrt{n}/\theta\right)\left(N_n/n\right)\left(k_n/b_n\right)$ with $k_n = \theta\sqrt{n}$ and $N_n/n \to 1$. Next, we show that $Var\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}\right) = O\left(\frac{k_n^2}{b_n^2}\right)$. We have that $Var\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}\right) \leq E\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}\right)^2$. Thus, using the c-r inequality, we can write

$$
\begin{aligned}
Var\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j}\right) &\leq \frac{nb_n^4 J_n}{\psi_2^4 k_n^4} \sum_{j=1}^{J_n-1} E\left(\Xi_{4j}^2\right) \\
&\leq K\frac{nJ_n}{\psi_2^4 k_n^4} J_n k_n^4 E\left(\bar{Y}_i^8\right), \quad (3.32)
\end{aligned}
$$

where the second inequality holds given Cauchy-Schwartz inequality and the fact that $E\left(\bar{Y}_i^8\right) = O\left(n^{-2}\right)$ uniformly in i. Thus, we can bound (3.32) by $K\frac{J_n^2}{n} = O\left(\frac{k_n^2}{b_n^2}\right)$ given that $J_n = N_n/b_n = \left(\sqrt{n}/\theta\right)\left(N_n/n\right)\left(k_n/b_n\right)$ with $k_n = \theta\sqrt{n}$ and $N_n/n \to 1$. Hence $\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \Xi_{4j} = O_P\left(\frac{k_n}{b_n}\right)$.

Proof of Lemma 3.1. Part a) Given the definition of V_n^*, we can write

$$V_n^* = V_{1n}^* - \frac{\sqrt{n} N_n b_n}{\left(N_n - b_n + 1\right)^2} V_{2n}^*,$$

where

$$\begin{aligned} V_{1n}^* &= \frac{1}{b_n} \sum_{t=0}^{b_n - 1} v_{1n,t}^*, \text{ with } v_{1n,t}^* \equiv \frac{\sqrt{n}}{\left(N_n - b_n + 1\right) N_n} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \left(\sum_{i=t+1}^{b_n + t} Z_{i+(j-1)b_n} \right)^2, \text{ and} \\ V_{2n}^* &= \frac{1}{b_n} \sum_{t=0}^{b_n - 1} v_{2n,t}^*, \text{ with } v_{2n,t}^* \equiv \frac{1}{N_n} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \sum_{i=t+1}^{b_n + t} Z_{i+(j-1)b_n}. \end{aligned}$$

We now proceed in two steps. In Step 1, we show that $v_{1n,t}^* \to^P V_p + \theta \left(p+1\right) \int_0^1 \left(\sigma_s^2 + \frac{\psi_1}{\theta^2 \psi_2} \omega^2\right)^2 ds$ uniformly in t. In Step 2, we show that $v_{2n,t}^* \to^P \int_0^1 \left(\sigma_s^2 + \frac{\psi_1}{\theta^2 \psi_2} \omega^2\right)^2 ds$, also uniformly in t. This together with the fact that $\frac{\sqrt{n} N_n b_n}{\left(N_n - b_n + 1\right)^2} \to \left(p+1\right) \theta$ as $n \to \infty$ when $b_n = \left(p+1\right) k_n$ and k_n satisfies Assumption 2 imply the result. Proof of Step 1. For $t = 0, \ldots, b_n - 1$ and $j = 1, \ldots, \left[\frac{N_n - t}{b_n}\right]$, let

$$\bar{B}_{j,t} \equiv \frac{1}{b_n} \sum_{i=1}^{b_n} \bar{Y}_{i-1+t+(j-1)b_n}^2 = \frac{k_n \psi_2}{N_n} \frac{1}{b_n} \sum_{i=1}^{b_n} Z_{i+t+(j-1)b_n},$$

where $Z_i \equiv \frac{N_n}{k_n} \frac{1}{\psi_2} \bar{Y}_{i-1}^2$ and note that the $\bar{B}_{j,t}$ are averages of non-overlapping blocks for given t. With this notation, we have that

$$v_{1n,t}^* = \frac{N_n^2}{\left(N_n - b_n + 1\right) N_n} \frac{\sqrt{n} b_n^2}{k_n^2 \psi_2^2} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \bar{B}_{j,t}^2,$$

where we can show that $\frac{N_n^2}{\left(N_n - b_n + 1\right) N_n} \to 1$ under the condition that $b_n = \left(p+1\right) k_n$. Using arguments similar to those used to prove Lemma 3.10.2, we can show that

$$\frac{\sqrt{n} b_n^2}{k_n^2 \psi_2^2} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \bar{B}_{j,t}^2 \to^P V_p + \theta \left(p+1\right) \int_0^1 \left(\sigma_s^2 + \frac{\psi_1}{\theta^2 \psi_2} \omega^2\right)^2 ds$$

uniformly in t. The proof of Step 2 relies on the consistency result in Theorem 1 of Christensen, Kinnebrock and Podolskij (2010). Indeed $v_{2n,t}^*$ is the main term in Jacod et al. (2009) pre-averaged realized volatility estimator without the bias corrected term, with starting point t. **Part b).** Follows directly from part a) of Lemma 2.3.1 when replacing σ_t by a constant for all t. **Part c).** Follows directly from part a) of Lemma 2.3.1.

Proof of Lemma 2.3.2 Part a). Given the definition of V_n^*, we can write

$$\begin{aligned} V_n^* &= Var^*\left(n^{1/4}PRV_n^*\right) = \frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\sum_{j=1}^{J_n-1}\left(\bar{B}_j-\bar{B}_{j+1}\right)^2 Var^*\left(\eta_j^2\right) \\ &= 2Var^*(\eta)\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\left(\sum_{j=1}^{J_n}\bar{B}_j^2-\sum_{j=1}^{J_n-1}\bar{B}_j\bar{B}_{j+1}\right) \\ &\quad -Var^*(\eta)\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\left(\bar{B}_1^2+\bar{B}_{J_n}^2\right). \end{aligned}$$

The result follows from Lemma 3.10.2 and the fact that $\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\left(\bar{B}_1^2+\bar{B}_{J_n}^2\right) = O_P\left(\frac{1}{\sqrt{n}}\right)$ given that $\bar{B}_j = O_P\left(1/\sqrt{n}\right)$ uniformly in j. **Part b).** Follows directly from Lemma 2.3.2.a) and the assumptions that $Var^*(\eta) = \frac{1}{2}$ that $p \to \infty$.

Proof of Theorem 2.3.1 For any fixed $p \geq 1$, let $S_n^* = n^{1/4}\left(PRV_n^* - E^*\left(PRV_n^*\right)\right) = \frac{b_n}{\psi_2 k_n}\sum_{j=1}^{J_n} z_j^*$, where $z_j^* = n^{1/4}\frac{b_n}{\psi_2 k_n}\left(\bar{B}_j{}^* - E^*\left(\bar{B}_j{}^*\right)\right)$. It follows that $E^*\left(\sum_{j=1}^{J_n} z_j^*\right) = 0$, and

$$V_n^* \equiv Var^*\left(\sum_{j=1}^{J_n} z_j^*\right) \xrightarrow{P} V_p + O_P\left(\frac{1}{p}\right) \equiv \tilde{V}_p.$$

Since $z_1^*\cdots, z_{J_n}^*$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta > 0$ and for some constant $C_p > 0$,

$$\sup_{x\in\Re}\left|P^*\left(S_n^*\leq x\right) - \Phi\left(x/\sqrt{\tilde{V}_p}\right)\right| \leq C_p\sum_{j=1}^{J_n}E^*\left|z_j^*\right|^{2+\delta},$$

which converges to zero in probability as $n \to \infty$. We have

$$\begin{aligned} \sum_{j=1}^{J_n}E^*\left|z_j^*\right|^{2+\delta} &= \sum_{j=1}^{J_n}E^*\left|n^{1/4}\frac{b_n}{\psi_2 k_n}\left(\bar{B}_j^*-E^*\left(\bar{B}_j^*\right)\right)\right|^{2+\delta} \\ &\leq 2n^{\frac{(2+\delta)}{4}}\left(\frac{b_n}{\psi_2 k_n}\right)^{2+\delta}\sum_{j=1}^{J_n}E^*\left|\bar{B}_j^*\right|^{2+\delta} \\ &\leq 2K_p n^{\frac{(2+\delta)}{4}}E^*\left|\eta_1\right|^{2+\delta}\sum_{j=1}^{J_n}\left|\bar{B}_j\right|^{2+\delta} = K_p O_p\left(n^{-\frac{\delta}{4}}\right) = o_p(1), \end{aligned}$$

since $E^*\left|\eta_j\right|^{2+\delta} \leq \Delta < \infty$, $\bar{B}_j = K_p O_p\left(\frac{1}{\sqrt{n}}\right)$, and $J_n \sim n^{1/2}$. It follows that $n^{1/4}\left(PRV_n^* - E^*\left(PRV_n^*\right)\right)$ $N(0, \tilde{V}_p)$ in probability, for any fixed $p \geq 1$. The result follows by using part $b)$ of Lemma 2.3.2 and letting $p \to \infty$.

3.11 Appendix for Chapter 3

3.11.1 Appendix E

This appendix is organized as follows. First, we details the cleaning we carried out on the data. Second, we report simulation results. Finally we report empirical results.

Data Cleaning

In line with Barndorff-Nielsen et al. (2008) we perform the following data cleaning steps:

(i) Delete entries outside the 9:30pm and 4pm time window.

(ii) Delete entries with a quote or transaction price equal to be zero.

(iii) Delete all entries with negative prices or quotes.

(iv) Delete all entries with negative spreads.

(v) Delete entries whenever the price is outside the interval $[bid - 2 * spread\ ;\ ask + 2 * spread]$.

(vi) Delete all entries with the spread greater or equal than 50 times the median spread of that day.

(vii) Delete all entries with the price greater or equal than 5 times the median mid-quote of that day.

(viii) Delete all entries with the mid-quote greater or equal than 10 times the mean absolute deviation from the local median mid-quote.

(ix) Delete all entries with the price greater or equal than 10 times the mean absolute deviation from the local median mid-quote.

We report in Table 3.7 below, the actual coverage rates for the feasible asymptotic theory approach and for our bootstrap methods. In Table 3.8 we summarize results using the optimal block size by minimizing confidence interval volatility. Table 3.9 provides some statistics of GE shares in August 2011.

Table 3.7. Coverage rates of nominal 95% CI for integrated volatility and integrated beta

	Integrated volatility									Integrated beta		
	SV1F				SV2F							
	Raw		Log		Raw		Log				Raw	
M	CLT	Boot	CLT	Boot	CLT	Boot	CLT	Boot		M	CLT	Boot
					$1/h = 12$							
1	85.44	98.49	90.08	97.86	80.38	96.62	86.17	96.24		2	83.66	95.88
2	85.56	97.31	90.31	96.80	80.43	94.70	86.27	94.73		3	87.63	95.03
3	85.71	96.46	90.84	96.08	80.34	93.77	85.89	93.70		4	89.14	94.83
4	85.88	96.20	90.97	95.93	80.34	92.88	85.52	92.89		6	90.67	94.49
12	86.11	94.84	91.27	94.87	77.66	88.89	81.65	86.97		12	90.44	93.63
					$1/h = 48$							
1	92.04	98.55	93.51	97.71	88.28	97.09	90.93	96.67		3	92.40	95.88
2	92.10	97.28	93.59	96.50	88.13	95.63	91.08	95.48		4	92.69	95.34
4	92.20	96.40	93.80	95.80	88.16	94.55	91.10	94.53		8	92.93	94.69
8	92.33	95.60	93.88	95.18	87.89	93.32	90.33	93.20		12	92.67	93.78
48	92.74	95.06	94.22	95.04	81.83	86.63	82.92	84.57		48	91.63	92.43
					$1/h = 96$							
1	93.35	97.94	94.09	97.10	90.20	97.06	92.10	96.66		3	92.62	95.57
2	93.43	96.78	93.99	96.06	90.37	95.84	92.24	95.67		4	93.13	95.00
4	93.47	95.78	94.03	95.61	90.46	94.70	92.09	94.83		8	93.83	94.84
8	93.50	95.26	94.09	93.32	90.07	93.81	91.75	94.01		12	93.77	94.57
96	93.42	94.80	94.35	94.87	81.93	84.61	82.79	83.60		96	91.94	92.35
					$1/h = 288$							
1	94.57	97.09	94.61	96.25	93.39	97.44	93.96	96.76		3	93.87	95.79
2	94.56	96.00	94.61	95.67	93.51	96.35	93.95	95.95		4	94.72	95.64
4	94.62	95.48	94.67	95.36	93.50	95.57	93.98	95.28		8	94.95	95.43
8	94.55	95.26	94.81	95.19	93.43	95.06	93.82	94.75		12	94.66	94.99
288	94.46	94.78	94.84	94.99	82.43	83.86	83.34	83.53		288	90.04	90.32
					$1/h = 576$							
1	94.53	96.12	94.75	95.84	94.19	96.96	94.49	96.52		3	93.94	95.62
2	94.57	95.53	94.68	95.41	94.17	96.23	94.52	95.78		4	94.46	95.40
4	94.74	95.15	94.70	95.16	94.32	95.59	94.56	95.45		8	94.58	94.87
8	94.67	95.08	94.72	94.96	94.22	95.38	94.46	95.16		12	94.53	94.88
576	94.58	94.85	94.76	94.92	82.01	82.37	82.05	82.32		576	87.07	87.07
					$1/h = 1152$							
1	95.06	96.06	95.16	95.70	94.51	96.52	94.47	95.95		3	94.78	95.93
2	95.13	95.68	95.20	95.65	94.53	95.79	94.47	95.42		4	94.92	95.48
4	95.05	95.49	95.20	95.31	94.42	95.21	94.50	95.11		8	94.88	95.13
8	95.15	95.47	95.18	95.20	94.39	95.03	94.47	94.85		12	94.95	94.87
1152	94.86	94.97	94.83	94.91	82.60	82.73	82.85	82.89		1152	81.68	81.62

Notes: CLT-intervals based on the Normal; Boot-intervals based on our proposed new local Gaussian bootstrap; M is the block size used to compute confidence intervals. 10,000 Monte Carlo trials with 999 bootstrap replications each.

TABLE 3.8. Coverage rates of nominal 95% intervals for integrated volatility and integrated beta using the optimal block size

	Integrated volatility																	
	SV1F									SV2F								
		Raw				Log					Raw				Log			
n	M^*	CLT	iidB	WB	Boot	CLT	iidB	WB	Boot	M^*	CLT	iidB	WB	Boot	CLT	iidB	WB	Boot
12	3.75	86.44	93.66	87.42	96.51	90.79	96.11	88.35	96.07	3.84	80.42	90.82	79.41	93.21	86.70	93.43	80.41	93.35
48	5.37	90.89	94.62	93.98	95.76	92.31	95.45	94.71	95.35	5.75	85.40	92.77	89.98	93.63	87.59	94.15	90.82	93.80
96	5.86	93.01	94.67	94.38	95.29	93.68	95.48	94.85	95.11	5.81	88.94	94.05	92.01	94.28	90.48	94.28	93.11	94.21
288	5.66	94.49	94.70	94.98	95.05	94.60	94.86	94.81	94.96	5.94	93.32	94.85	94.18	95.02	93.62	94.99	94.36	94.90
576	5.82	94.50	94.54	94.51	94.95	94.58	94.65	94.66	94.98	6.06	94.08	94.89	94.47	95.19	94.38	94.90	94.81	95.03
1152	6.01	95.05	94.87	95.13	95.04	95.15	94.85	95.14	95.02	6.25	94.41	94.83	94.56	94.92	94.38	94.92	94.76	94.83

	Integrated beta			
n	M^*	CLT	PairsB	Boot
12	3.64	88.49	93.59	95.17
48	4.97	92.86	93.96	94.84
96	5.60	93.70	94.98	94.62
288	5.76	94.55	94.75	94.98
576	5.78	94.26	94.72	94.64
1152	5.92	94.77	94.67	94.78

Notes: CLT-intervals based on the Normal; iidB-intervals based on the i.i.d. bootstrap of Gonçalves and Meddahi (2009); WB-wild bootstrap based on Proposition 4.5 of Gonçalves and Meddahi (2009); Boot-intervals based on our proposed new local Gaussian bootstrap; PairsB-intervals based on the pairs bootstrap of Dovonon Gonçalves and Meddahi (2013); M^* is the optimal block size selected by using the Minimum Volatility method. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 3.9. Summary statistics

Days	Trans	n	S
1 Aug	11303	96	118
2 Aug	13873	96	145
3 Aug	13205	96	138
4 Aug	16443	96	172
5 Aug	16212	96	169
8 Aug	18107	96	189
9 Aug	18184	96	190
10 Aug	15826	96	165
11 Aug	15148	96	158
12 Aug	12432	96	130
15 Aug	12042	96	126
16 Aug	10128	96	106
17 Aug	9104	96	95
18 Aug	15102	96	158
19 Aug	11468	96	120
22 Aug	10236	96	107
23 Aug	11518	96	120
24 Aug	10429	96	109
25 Aug	9794	96	102
26 Aug	9007	96	94
29 Aug	10721	96	112
30 Aug	9131	96	96
31 Aug	10724	96	112

"Trans" denotes the number of transactions, n the sample size used to compute the realized volatility, and sampling of every S'th transaction price, so the period over which returns are calculated is roughly 15 minutes.

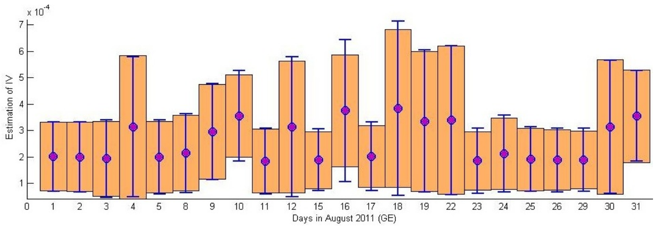

Figure 3.3. 95% Confidence Intervals (CI's) for the daily $\overline{\sigma^2}$, for each regular exchange opening days in August 2011, calculated using the asymptotic theory of Mykland and Zhang (CI's with bars), and the new wild bootstrap method (CI's with lines). The realized volatility estimator is the middle of all CI's by construction. Days on the x-axis.

3.11.2 Appendix F

This appendix concerns only the case where $d = 1$ (i.e. when the parameter of interest is the integrated volatility). We organized this appendix as follows. First, we introduce some notation. Second, we state Lemmas 3.11.1 and 3.11.2, Theorems 3.11.1 and 3.11.2 and their proofs useful for proofs for the theorem 3.3.1 and proposition 3.5.1 presented in the main text. These results are used to obtain the formal Edgeworth expansions through order $O(h)$ for realized volatility. Finally, we prove the Theorem 3.3.1 and the Propositions 3.5.1.

Notation

To make for greater comparability, and in order to use some existing results, we have kept the notation from Gonçalves and Meddahi (2009) whenever possible. We introduce some notation, recall that, for any $q > 0$, $\overline{\sigma^q} \equiv \int_0^1 \sigma_u^q du$, and let $\overline{\sigma_{h,M}^q} \equiv Mh \sum_{j=1}^{1/Mh} \left(\frac{\sigma_{j,M}^2}{Mh}\right)^{q/2}$, where $\sigma_{j,M}^2 \equiv \int_{(j-1)Mh}^{jMh} \sigma_u^2 du = Mh\sigma_{(j-1)Mh}^2 > 0$. We let $\sigma_{q,p} \equiv \frac{\overline{\sigma^q}}{\left(\overline{\sigma^p}\right)^{q/p}}$, when $\overline{\sigma^q}$ is replaced with $\overline{\sigma_{h,M}^q}$ we write $\sigma_{q,p,h,M}$, and $R_q \equiv Mh \sum_{j=1}^{1/Mh} \left(\frac{RV_{j,M}}{Mh}\right)^{q/2}$, where $RV_{j,M} = \sum_{i=1}^{M} y_{i+(j-1)M}^2$. We also let $R_{q,p} \equiv \frac{R_q}{(R_p)^{q/p}}$. Recall that $c_{M,q} \equiv E\left(\left(\frac{\chi_M^2}{M}\right)^{q/2}\right)$ with χ_M^2 the standard χ^2 distribution with M degrees of freedom. Note that $c_{M,2} = 1$, $c_{M,4} = \frac{M+2}{M}$, $c_{M,6} = \frac{(M+2)(M+4)}{M^2}$ and $c_{M,8} = \frac{(M+2)(M+4)(M+6)}{M^3}$. It follows by using the definition of $c_{M,q}$ gives in equation (3.6) and this property of the Gamma function, for all $x > 0$, $\Gamma(x+1) = x\Gamma(x)$. We follow Gonçalves and Meddahi (2009) and we write

$$T_{h,M} = S_{h,M}\left(\frac{\hat{V}}{V_{h,M}}\right)^{-1/2} = S_{h,M}\left(1 + \sqrt{h}U_{h,M}\right)^{-1/2},$$

where

$$S_{h,M} = \frac{\sqrt{h^{-1}}\left(R_2 - c_{M,2}\overline{\sigma^2}\right)}{\sqrt{V_{h,M}}} \text{ and } U_{h,M} \equiv \frac{\sqrt{h^{-1}}\left(\hat{V} - V_{h,M}\right)}{V_{h,M}},$$

and $V_{h,M} = Var\left(\sqrt{h^{-1}}R_2\right) = M\left(c_{M,4} - c_{M,2}^2\right)\overline{\sigma_{h,M}^4}$. The proof of Lemma 3.11.1 below relies heavily on the fact that, for any $q > 0$, $\left|RV_{j,M}\right|^{q/2} - c_{M,q}\left|\sigma_{j,M}^2\right|^{q/2}$ are conditionally on σ independent with zero mean since $RV_{j,M} = \sigma_{j,M}^2 \frac{\chi_{j,M}^2}{M}$ where $\frac{\chi_{j,M}^2}{M} \equiv \frac{\sum_{i=1}^{M} \eta_{(j-1)M+i}^2}{M}$ and $\eta_i \sim i.i.d.\ N(0,1)$. We rewrite $R_2 - c_{M,2}\overline{\sigma^2}$ and $\hat{V} - V_{h,M}$ as follows

$$R_2 - c_{M,2}\overline{\sigma^2} = \sum_{j=1}^{1/Mh}\left(RV_{j,M} - c_{M,2}\sigma_{j,M}^2\right),$$

$$\hat{V} - V_{h,M} = M\left(\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\right)(Mh)^{-1}\sum_{j=1}^{1/Mh}\left(RV_{j,M}^2 - c_{M,4}\sigma_{j,M}^4\right).$$

Similarly for the bootstrap, we let $T_{h,M}^* = S_{h,M}^*\left(1+\sqrt{h}U_{h,M}^*\right)^{-1/2}$, where $S_{h,M}^* = \frac{\sqrt{h^{-1}}\left(R_2^* - c_{M,2}R_2\right)}{\sqrt{V^*}}$, $U_{h,M}^* \equiv \frac{\sqrt{h^{-1}}\left(\hat{V}^* - V^*\right)}{V^*}$ and $V^* = Var^*\left(n^{1/2}R_2^*\right)$.

Finally, note that throughout we will use $\sum_{i\neq j\neq k} = \sum_{i\neq j, i\neq k, j\neq k}$, to denote a sum where all indices differ, e.g.

Lemma 3.11.1. *Suppose (1) holds, conditionally on σ, and under Q_h for any $q>0$, and any $M \geq 1$ such that $M \approx cn^{-\alpha}$ with $\alpha \in [0, 1/2)$, we have*

a1) $E\left(\left|RV_{j,M}\right|^{q/2}\right) = c_{M,q}\left|\sigma_{j,M}^2\right|^{q/2}$,

a2) $V_{h,M} \equiv Var\left(\sqrt{h^{-1}}R_2\right) = M\left(c_{M,4} - c_{M,2}^2\right)\overline{\sigma_{h,M}^4}$,

a3) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^3\right] = (Mh)^2\left(c_{M,6} - 3c_{M,2}c_{M,4} + 2c_{M,2}^3\right)\overline{\sigma_{h,M}^6}$,

a4)

$$\begin{aligned} E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^4\right] &= 3(Mh)^2\left(c_{M,4} - c_{M,2}^2\right)^2\left(\overline{\sigma_{h,M}^4}\right)^2 \\ &+ (Mh)^3\left(c_{M,8} - 3c_{M,2}c_{M,6} + 12c_{M,2}^2c_{M,4} - 6c_{M,2}^4 - 3c_{M,4}^2\right)\overline{\sigma_{h,M}^8}, \end{aligned}$$

a5)

$$E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)\left(\hat{V} - V_{h,M}\right)\right] = M\left(\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\right)(Mh)\left(c_{M,6} - c_{M,2}c_{M,4}\right)\overline{\sigma_{h,M}^6},$$

a6) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^2\left(\hat{V} - V_{h,M}\right)\right] = M\left(\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\right)(Mh)^2\left[\begin{array}{c} c_{M,8} - c_{M,4}^2 \\ -2c_{M,2}c_{M,6} + c_{M,2}^2c_{M,4}\end{array}\right]\overline{\sigma_{h,M}^8}$,

a7) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^3\left(\hat{V} - V_{h,M}\right)\right] = 3M(Mh)^2\frac{\left(c_{M,4} - c_{M,2}^2\right)^2\left(c_{M,6} - c_{M,2}c_{M,4}\right)}{c_{M,4}}\left(\overline{\sigma_{h,M}^4}\right)\left(\overline{\sigma_{h,M}^4}\right)$
$+ 384h^3\overline{\sigma_{h,M}^{10}}$
$= 3M(Mh)^2\frac{\left(c_{M,4} - c_{M,2}^2\right)^2\left(c_{M,6} - c_{M,2}c_{M,4}\right)}{c_{M,4}}\left(\overline{\sigma_{h,M}^4}\right)\left(\overline{\sigma_{h,M}^4}\right)$
$+ O\left(h^3\right)$ *as* $h \to 0$,

a8) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^4\left(\hat{V} - V_{h,M}\right)\right] =$
$(Mh)^3 M\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\left[\begin{array}{c} 4\left(c_{M,6} - 3c_{M,2}c_{M,4} + 2c_{M,2}^3\right)\left(c_{M,6} - c_{M,2}c_{M,4}\right)\left(\overline{\sigma_{h,M}^6}\right)^2 \\ +6\left(c_{M,8} - c_{M,4}^2 - 2c_{M,2}c_{M,6} + 2c_{M,2}^2c_{M,4}\right)\left(c_{M,4} - c_{M,2}^2\right)\left(\overline{\sigma_{h,M}^4}\right)\left(\overline{\sigma_{h,M}^8}\right)\end{array}\right]$
$+ O\left(h^4\right)$ *as* $h \to 0$,

a9) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)\left(\hat{V} - V_{h,M}\right)^2\right] = O\left(h^2\right)$ *as* $h \to 0$,

a10) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^2\left(\hat{V} - V_{h,M}\right)^2\right] =$

$$(Mh)^2 M\frac{\left(c_{M,4}-c_{M,2}^2\right)^2}{c_{M,4}^2}\left[\begin{array}{c}\left(c_{M,4}-c_{M,2}^2\right)\left(c_{M,8}-c_{M,4}^2\right)\left(\overline{\sigma_{h,M}^4}\right)\left(\overline{\sigma_{h,M}^8}\right)\\ +2\left(c_{M,6}-c_{M,2}c_{M,4}\right)^2\left(\overline{\sigma_{h,M}^6}\right)^2\end{array}\right]$$

$+ O\left(h^3\right)$ *as* $h \to 0$,

a11) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^3\left(\hat{V} - V_{h,M}\right)^2\right] = O\left(h^3\right)$ *as* $h \to 0$,

a12) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^4\left(\hat{V} - V_{h,M}\right)^2\right] =$

$$(Mh)^3 M\frac{\left(c_{M,4}-c_{M,2}^2\right)^2}{c_{M,4}^2}\left[\begin{array}{c}3\left(c_{M,4}-c_{M,2}^2\right)^2\left(c_{M,8}-c_{M,4}^2\right)\left(\overline{\sigma_{h,M}^4}\right)^2\left(\overline{\sigma_{h,M}^8}\right)\\ +12\left(c_{M,4}-c_{M,2}^2\right)^2\left(c_{M,6}-c_{M,2}c_{M,4}\right)^2\left(\overline{\sigma_{h,M}^6}\right)^2\left(\overline{\sigma_{h,M}^4}\right)\end{array}\right]$$

$+ O\left(h^4\right)$ *as* $h \to 0$.

Lemma 3.11.2. *Suppose (1) holds, conditionally on* σ, *and under* Q_h *for any* $M \geq 1$ *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in (1/2, 1]$*), we have*

$$\begin{aligned}
E\left(S_{h,M}\right) &= 0,\\
E\left(S_{h,M}^2\right) &= 1,\\
E\left(S_{h,M}^3\right) &= \sqrt{h}B_{1,M}\sigma_{6,4,h,M},\\
E\left(S_{h,M}^4\right) &= 3 + hB_{2,M}\sigma_{8,4,h,M},\\
E\left(S_{h,M}U_{h,M}\right) &= A_{1,M}\sigma_{6,4,h,M},\\
E\left(S_{h,M}^2U_{h,M}\right) &= \sqrt{h}A_{2,M}\sigma_{8,4,h,M},
\end{aligned}$$

and as $h \to 0$ *we have,*

$$\begin{aligned}
E\left(S_{h,M}^3U_{h,M}\right) &= A_{3,M}\sigma_{6,4,h,M} + O\left(h\right),\\
E\left(S_{h,M}^4U_{h,M}\right) &= \sqrt{h}[D_{1,M}\sigma_{8,4,h,M} + D_{2,M}\sigma_{6,4,h,M}^2] + O\left(h^{3/2}\right),\\
E\left(S_{h,M}U_{h,M}^2\right) &= O\left(h^{1/2}\right),\\
E\left(S_{h,M}^3U_{h,M}^2\right) &= O\left(h^{1/2}\right),\\
E\left(S_{h,M}^2U_{h,M}^2\right) &= [C_{1,M}\sigma_{8,4,h,M} + C_{2,M}\sigma_{6,4,h,M}^2] + O\left(h\right),\\
E\left(S_{h,M}^4U_{h,M}^2\right) &= [E_{1,M}\sigma_{8,4,h,M} + E_{2,M}\sigma_{6,4,h,M}^2] + O\left(h\right),
\end{aligned}$$

where,

$$
\begin{aligned}
A_{1,M} &= \frac{1}{\sqrt{M}} \frac{c_{M,6} - c_{M,2}c_{M,4}}{c_{M,4}\left(c_{M,4} - c_{M,2}^2\right)^{1/2}} = \frac{2\sqrt{2}}{M}, \\
B_{1,M} &= \sqrt{M}\frac{\left(c_{M,6} - 3c_{M,2}c_{M,4} + 2c_{M,2}^3\right)}{\left(c_{M,4} - c_{M,2}^2\right)^{3/2}} = 2\sqrt{2}, \\
A_{2,M} &= \frac{c_{M,8} - c_{M,4}^2 - 2c_{M,2}c_{M,6} + 2c_{M,2}^2 c_{M,4}}{c_{M,4}\left(c_{M,4} - c_{M,2}^2\right)} = \frac{12}{M}, \\
B_{2,M} &= M\frac{c_{M,8} - 4c_{M,2}c_{M,6} + 12c_{M,2}^2 c_{M,4} - 6c_{M,2}^4 - 3c_{M,4}^2}{\left(c_{M,4} - c_{M,2}^2\right)^2} = 12, \\
C_{1,M} &= \frac{c_{M,8} - c_{M,4}^2}{c_{M,4}^2 M} = \frac{24 + 8M}{M^2\left(2 + M\right)},
\end{aligned}
$$

with $A_{3,M} = 3A_{1,M}$, $C_{2,M} = 2A_{1,M}^2$, $D_{1,M} = 6A_{2,M}$, $D_{2,M} = 4A_{1,M}B_{1,M}$, $E_{1,M} = 3C_{1,M}$ *and* $E_{2,M} = 12A_{1,M}^2$.

Remark 4 The bootstrap analogue of Lemma 3.11.1 replaces $RV_{j,M}$ with $RV_{j,M}^*$, $\sigma_{j,M}^2$ with $RV_{j,M}$, and $\overline{\sigma_{h,M}^q}$ with R_q, the bootstrap analogue of Lemma 3.11.2 replaces $\sigma_{q,p,h,M}$ with $R_{q,p}$.

Theorem 3.11.1. ***(Cumulants of** $T_{h,M}$**)*** *Consider DGP (3.1) and suppose assumption H holds. Then for any* $q > 0$*, and any* $M \geq 1$ *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$, $\overline{\sigma_{h,M}^q} = \overline{\sigma_h^q}$ *and* $\overline{\sigma^q} - \overline{\sigma_{h,M}^q} = o_P(\sqrt{h})$*, conditionally on* σ *and under* Q_h*, it follows that as* $h \to 0$*,*

$$
\begin{aligned}
\kappa_1\left(T_{h,M}\right) &= \sqrt{h}\kappa_1 + o(h) \text{ with } \kappa_1 = -\frac{A_{1,M}}{2}\sigma_{6,4}; \\
\kappa_2\left(T_{h,M}\right) &= 1 + h\kappa_2 + o(h) \text{ with } \kappa_2 = \left(C_{1,M} - A_{2,M}\right)\sigma_{8,4} + \frac{7}{4}A_{1,M}^2\sigma_{6,4}^2; \\
\kappa_3\left(T_{h,M}\right) &= \sqrt{h}\kappa_3 + o(h) \text{ with } \kappa_3 = \left(B_{1,M} - 3A_{1,M}\right)\sigma_{6,4}; \\
\kappa_4\left(T_{h,M}\right) &= h\kappa_4 + o(h) \text{ with } \kappa_4 = \left(B_{2,M} + 3C_{1,M} - 6A_{2,M}^2\right)\sigma_{8,4} + \left(18{A_{1,M}}^2 - 6A_{1,M}B_{1,M}\right)\sigma_{6,4}^2.
\end{aligned}
$$

Note that $A_{1,M}$, $A_{2,M}$, $B_{1,M}$, $B_{2,M}$*, and* $C_{1,M}$*, are as in Lemma 3.11.2, and* $A_{3,M} = 3A_{1,M}$, $C_{2,M} = 2A_{1,M}^2$, $D_{1,M} = 6A_{2,M}$, $D_{2,M} = 4A_{1,M}B_{1,M}$, $E_{1,M} = 3C_{1,M}$ *and* $E_{2,M} = 12A_{1,M}^2$.

Theorem 3.11.2. ***(Bootstrap Cumulants of** $T_{h,M}^*$**)*** *Consider DGP (3.1) and suppose (3.5) holds. Let* $M \geq 1$ *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$*, under assumption H,* *conditionally on σ, it follows that as $h \to 0$*

$$
\begin{aligned}
\kappa_1^*\left(T_{h,M}^*\right) &= \sqrt{h}\kappa_{1,h,M}^* + o(h) \text{ with } \kappa_{1,h,M}^* = -\frac{A_{1,M}}{2} R_{6,4};\\
\kappa_2^*\left(T_{h,M}^*\right) &= 1 + h\kappa_{2,h,M}^* + o(h) \text{ with } \kappa_{2,h,M}^* = \left(C_{1,M} - A_{2,M}\right) R_{8,4} + \frac{7}{4} A_{1,M}^2 R_{6,4}^2;\\
\kappa_3^*\left(T_{h,M}^*\right) &= \sqrt{h}\kappa_{3,h,M}^* + o(h) \text{ with } \kappa_{3,h,M}^* = \left(B_{1,M} - 3A_{1,M}\right) R_{6,4};\\
\kappa_4^*\left(T_{h,M}^*\right) &= h\kappa_{4,h,M}^* + o(h) \text{ with } \kappa_{4,h,M}^* = \left(B_{2,M} + 3C_{1,M} - 6A_{2,M}^2\right) R_{8,4} + \left(18{A_{1,M}}^2 - 6A_{1,M}B_{1,M}\right)
\end{aligned}
$$

Note that $A_{1,M}$, $A_{2,M}$, B_1, M, $B_{2,M}$, and C_1, M, are as in Lemma 3.11.2, and $A_{3,M} = 3A_{1,M}$, $C_{2,M} = 2A_{1,M}^2$, $D_{1,M} = 6A_{2,M}$, $D_{2,M} = 4A_{1,M}B_{1,M}$, $E_{1,M} = 3C_{1,M}$ and $E_{2,M} = 12A_{1,M}^2$.

Proof of Lemma 3.11.1 *a*1) follows from $RV_{j,M} = \sum_{i=1}^{M} y_{(j-1)M+i}^2 = \frac{\chi_{j,M}^2}{M}\sigma_{j,M}^2$, where $\frac{\chi_{j,M}^2}{M} \equiv \frac{\sum_{i=1}^{M}\eta_{(j-1)M+i}^2}{M}$ and $\eta_i \sim i.i.d.\ N(0,1)$. For *a*2) note that $R_2 = \sum_{j=1}^{1/Mh} RV_{j,M}$, where $RV_{j,M}$ is conditional on σ independent with $Var\left(RV_{j,M}\right) = \left(c_{M,4} - c_{M,2}^2\right)\sigma_{j,M}^4$, with $\sigma_{j,M}^4 = \left(\sigma_{j,M}^2\right)^2$. It follows that,

$$
\begin{aligned}
Var\left(\sqrt{h^{-1}}R_2\right) &= h^{-1}\sum_{j=1}^{1/Mh} Var\left(RV_{j,M}\right)\\
&= h^{-1}\left(c_{M,4} - c_{M,2}^2\right)\sum_{j=1}^{1/Mh}\sigma_{j,M}^4\\
&= M\left(c_{M,4} - c_{M,2}^2\right)\overline{\sigma_{h,M}^4}.
\end{aligned}
$$

To prove the remaining results we follow the same structure of proofs as Gonçalves and Meddahi (2009). Here $c_{M,4}$ plays the role of $\mu_q = E\left(|\eta|^q\right)$, where $\eta \sim i.i.d.\ N(0,1)$ in Gonçalves and Meddahi (2009) and $RV_{j,M}$ plays the role of r_i^2 in Gonçalves and Meddahi (2009).

Proof of Lemma 3.11.2 Results follow directly from the definition of $S_{h,M}$, $U_{h,M}$ and Lemma 3.11.1.

Proof of Theorem 3.11.1 The first four cumulants of $T_{h,M}$ are given by (e.g., Hall, 1992, p.42):

$$\begin{aligned}
\kappa_1(T_{h,M}) &= E(T_{h,M}), \\
\kappa_2(T_{h,M}) &= E\left(T_{h,M}^2\right) - (E(T_{h,M}))^2, \\
\kappa_3(T_{h,M}) &= E\left(T_{h,M}^3\right) - 3E\left(T_{h,M}^2\right)E(T_{h,M}) + 2(E(T_{h,M}))^3, \\
\kappa_4(T_{h,M}) &= E\left(T_{h,M}^4\right) - 4E\left(T_{h,M}^3\right)E(T_{h,M}) - 3(E\left(T_{h,M}^2\right))^2 + 12E\left(T_{h,M}^2\right)(E(T_{h,M}))^2 \\
&- 6(E(T_{h,M}))^4.
\end{aligned}$$

Our goal is to identify the terms of order up to $O(h)$ in the asymptotic expansions of these four cumulants. We will first provide asymptotic expansions through order $O(h)$ for the first four moments of $T_{h,M}$ by using a Taylor expansion. For a fixed value k, a second-order Taylor expansion of $f(x) = (1+x)^{-k/2}$ around 0 yields $f(x) = 1 - \frac{k}{2}x + \frac{k}{4}(\frac{k}{2}+1)x^2 + O(x^3)$. We have that for any fixed integer k,

$$\begin{aligned}
T_{h,M}^k &= S_{h,M}^k\left(1+\sqrt{h}U_{h,M}\right)^{-k/2} + O(h^{3/2}), \\
&= S_{h,M}^k - \frac{k}{2}\sqrt{h}S_{h,M}^k U_{h,M} + \frac{k}{4}(\frac{k}{2}+1)hS_{h,M}U_{h,M}^2 + O(h^{3/2}).
\end{aligned}$$

For $k = 1, \cdots, 4$, the moments of $T_{h,M}^k$ up to order $O(h^{3/2})$ are given by

$$\begin{aligned}
E(T_{h,M}) &= 0 - \frac{\sqrt{h}}{2}E(S_{h,M}U_{h,M}) + \frac{3}{8}hE\left(S_{h,M}U_{h,M}^2\right) \\
E\left(T_{h,M}^2\right) &= 1 - \sqrt{h}E\left(S_{h,M}^2U_{h,M}\right) + hE\left(S_{h,M}^2U_{h,M}^2\right) \\
E\left(T_{h,M}^3\right) &= E\left(S_{h,M}^3\right) - \sqrt{h}\frac{3}{2}E\left(S_{h,M}^3U_{h,M}\right) + \frac{15}{8}hE\left(S_{h,M}^3U_{h,M}^2\right) \\
E\left(T_{h,M}^4\right) &= E\left(S_{h,M}^4\right) - 2\sqrt{h}E\left(S_{h,M}^4U_{h,M}\right) + 3hE\left(S_{h,M}^4U_{h,M}^2\right).
\end{aligned}$$

where we used $E(S_{h,M}) = 0$, and $E\left(S_{h,M}^2\right) = 1$. By Lemma 3.11.2 in Appendix B, we have that

$$\begin{aligned}
E(T_{h,M}) &= \sqrt{h}\left(-\frac{A_{1,M}}{2}\sigma_{6,4,h}\right) + O(h^{3/2}), \\
E\left(T_{h,M}^2\right) &= 1 + \sqrt{h}\left((C_{1,M} - A_{2,M})\sigma_{8,4} + C_{2,M}\sigma_{6,4,h}^2\right) + O(h^2) \\
E\left(T_{h,M}^3\right) &= \sqrt{h}\left(\left(B_{1,M} - \frac{3}{2}A_{3,M}\right)\sigma_{6,4,h}\right) + O(h^{3/2}) \\
E\left(T_{h,M}^4\right) &= 3 + h\left((B_{2,M} - 2D_{1,M} + 3E_{1,M})\sigma_{8,4,h} + (3E_{2,M} - 2D_{2,M})\sigma_{6,4,h}^2\right) + O(h^2).
\end{aligned}$$

Thus $\kappa_1\left(T_{h,M}\right) = \sqrt{h}\left(-\frac{A_{1,M}}{2}\sigma_{6,4,h}\right) + O(h^{3/2}) = \sqrt{h}\left(-\frac{A_{1,M}}{2}\sigma_{6,4}\right) + o(h^{3/2})$, since under Assumption H, BNS (2004) showed that $\overline{\sigma^q} - h^{1-q/2}\sum_{s=1}^{1/h}\left(\int_{(s-1)h}^{sh}\sigma_u^2 du\right)^{q/2} = o_P(\sqrt{h})$. Next we show that under (3.4) and given the defininion of $\overline{\sigma_{h,M}^q}$, we have $\overline{\sigma_{h,M}^q} = h^{1-q/2}\sum_{s=1}^{1/h}\left(\int_{(s-1)h}^{sh}\sigma_u^2 du\right)^{q/2}$. Note that, for any positive integer M, given the definitions of $\overline{\sigma_{h,M}^q}$ and $\sigma_{j,M}^2$, we can write

$$\begin{aligned}
\overline{\sigma_{h,M}^q} &= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}\left(\sigma_{j,M}^2\right)^{q/2} \\
&= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}\left(\int_{(j-1)Mh}^{jMh}\sigma_u^2 du\right)^{q/2},
\end{aligned}$$

using the fact that under Q_h, we have $\sigma_{j,M}^2 \equiv \int_{(j-1)Mh}^{jMh}\sigma_u^2 du = Mh\sigma_{(j-1)Mh}^2 > 0$, it follows that

$$\begin{aligned}
\overline{\sigma_{h,M}^q} &= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}\left(Mh\sigma_{(j-1)Mh}^2\right)^{q/2} \\
&= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}M^{q/2}\left(h\sigma_{(j-1)Mh}^2\right)^{q/2} \\
&= h^{1-q/2}\sum_{j=1}^{1/Mh}M\left(h\sigma_{(j-1)Mh}^2\right)^{q/2} \\
&= h^{1-q/2}\sum_{j=1}^{1/Mh}\sum_{i=1}^{M}\left(h\sigma_{(j-1)Mh}^2\right)^{q/2} \\
&= h^{1-q/2}\sum_{j=1}^{1/Mh}\sum_{i=1}^{M}\left(\int_{((j-1)M+i-1)h}^{((j-1)M+i)h}\sigma_u^2 du\right)^{q/2} \\
&= h^{1-q/2}\sum_{s=1}^{1/h}\left(\int_{(s-1)h}^{sh}\sigma_u^2 du\right)^{q/2}.
\end{aligned}$$

Thus $\overline{\sigma_{h,M}^q} = \overline{\sigma_h^q}$, this proves the first result. The remaining results follow similarly.

Proof of Theorem 3.11.2 See the proof of Theorem 3.11.1 and Remark 4.

Proofs of Theorem 3.3.1, and Proposition 3.5.1.

Proof of Theorem 3.3.1 Given that $T_{h,M} \xrightarrow{d} N(0,1)$, it suffices that $T_{h,M}^{*} \xrightarrow{d^*} N(0,1)$ in probability under Q_h. Let

$$H_{h,M}^{*} = \frac{\sqrt{h^{-1}}\left(R_2^{*} - E^{*}\left(R_2^{*}\right)\right)}{\sqrt{V^{*}}},$$

and note that

$$T_{h,M}^{*} = H_{h,M}^{*}\sqrt{\frac{V^{*}}{\hat{V}^{*}}}.$$

The proof contains two steps.
Step 1 We show that $H_{h,M}^{*} \xrightarrow{d^*} N(0,1)$ in probability under Q_h.
Step 2 We show that $\hat{V}^{*} \xrightarrow{P^*} V^{*}$ in probability under Q_h.
For step 1, we can write

$$H_{h,M}^{*} = \sum_{j=1}^{1/Mh} z_j^{*},$$

where

$$z_j^{*} = \frac{\sqrt{h^{-1}}\left(RV_{j,M}^{*} - E^{*}\left(RV_{j,M}^{*}\right)\right)}{\sqrt{V^{*}}}$$

with $E^{*}\left(\sum_{j=1}^{1/Mh} z_j^{*}\right) = 0$, and $Var^{*}\left(\sum_{j=1}^{1/Mh} z_j^{*}\right) = 1$.

Since $z_1^{*}, \ldots, z_{1/Mh}^{*}$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta > 0$ and for some constant $C > 0$ (which changes from line to line),

$$\sup_{x \in \Re}\left|P^{*}\left(H_{h,M}^{*} \leq x\right) - \Phi\left(x\right)\right| \leq C \sum_{j=1}^{1/Mh} E^{*}\left|z_j^{*}\right|^{2+\delta},$$

which converges to zero in probability for any $M \geq 1$ such that $M \approx ch^{-\alpha}$ with $\alpha \in [0, 1/2)$, as $h \to 0$. Indeed, we have that

$$\begin{aligned}
\sum_{j=1}^{1/Mh} E^{*}\left|z_j^{*}\right|^{2+\delta} &= \sum_{j=1}^{1/Mh} E^{*}\left|\frac{\sqrt{h^{-1}}\left(RV_{j,M}^{*} - E^{*}\left(RV_{j,M}^{*}\right)\right)}{\sqrt{V^{*}}}\right|^{2+\delta} \\
&\leq 2V^{*-\frac{(2+\delta)}{2}} h^{\frac{-(2+\delta)}{2}} \sum_{j=1}^{1/Mh} E^{*}\left|RV_{j,M}^{*}\right|^{2+\delta} \\
&= 2V^{*-\frac{(2+\delta)}{2}} h^{\frac{-(2+\delta)}{2}} E^{*}\left|\frac{\sum_{i=1}^{M}\eta_{(j-1)M+i}^{2}}{M}\right|^{2+\delta} \sum_{j=1}^{1/Mh}\left|RV_{j,M}\right|^{2+\delta},
\end{aligned}$$

where the inequality follows from the C_r and the Jensen inequalities. Then, given the definitions of $c_{M,2(2+\delta)}$ and $R_{2(2+\delta)}$, we can write

$$\begin{aligned}
\sum_{j=1}^{1/Mh} E^* \left| z_j^* \right|^{2+\delta} &\leq 2V^{*-\frac{(2+\delta)}{2}} c_{M,2(2+\delta)} M^{1+\delta} h^{\frac{\delta}{2}} R_{2(2+\delta)} \\
&\leq CV^{*-\frac{(2+\delta)}{2}} c_{M,2(2+\delta)}^2 h^{\frac{\delta}{2}-\alpha(1+\delta)} \frac{1}{c_{M,2(2+\delta)}} R_{2(2+\delta)} \\
&= O_p\left(h^{\frac{\delta}{2}-\alpha(1+\delta)} c_{M,2(2+\delta)}^2 \right) \\
&= o_p(1).
\end{aligned}$$

Note that for any $\delta > 0$ and $\alpha \in [0, 1/2)$ we have $\frac{\delta}{2} - \alpha(1+\delta) > 0$. Results follow since as $h \to 0$, $V^* \xrightarrow{P} 2\overline{\sigma^4} > 0$, and we have $\frac{1}{c_{M,2(2+\delta)}} R_{2(2+\delta)} \xrightarrow{P} \overline{\sigma^{2(2+\delta)}} = O(1)$, and $c_{M,2(2+\delta)} \to 1$.

For step 2, we show that $Bias^*\left(\widehat{V^*}\right) \xrightarrow{Q_h} 0$ and $Var^*\left(\widehat{V^*}\right) \xrightarrow{Q_h} 0$.
We have that

$$\begin{aligned}
Bias^*\left(\widehat{V^*}\right) &= E^*\left(\widehat{V^*}\right) - V^* \\
&= M\left(\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\right)(Mh)^{-1} \sum_{j=1}^{1/Mh} E^*\left(RV_{j,M}^{2*} - c_{M,4} RV_{j,M}^2 \right) \\
&= 0,
\end{aligned}$$

we also have,

$$\begin{aligned}
Var^*\left(\widehat{V^*}\right) &= E^*\left(\widehat{V^*} - V^*\right)^2 - \left(E^*\left(\widehat{V^*} - V^*\right)\right)^2 \\
&= M^2\left(\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\right)^2 (Mh)^{-2} E^*\left(\sum_{j=1}^{1/Mh}\left(RV_{j,M}^{2*} - c_{M,4} RV_{j,M}^2 \right)\right)^2 \\
&= M^2\left(\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\right)^2 (Mh)^{-2} \sum_{j=1}^{1/Mh} RV_{j,M}^4 E^*\left(\left(\frac{\chi_{j,M}^2}{M}\right)^2 - c_{M,4}\right)^2,
\end{aligned}$$

then, given the definitions of $c_{M,2}, c_{M,4}$, $c_{M,8}$ and R_8, we can write

$$\begin{aligned}
Var^*\left(\widehat{V^*}\right) &= M^2\left(\frac{c_{M,4}-c_{M,2}^2}{c_{M,4}}\right)^2 (Mh)^{-2}\left(c_{M,8}-c_{M,4}^2\right)\sum_{j=1}^{1/Mh} RV_{j,M}^4 \\
&= M^2\left(\frac{c_{M,4}-c_{M,2}^2}{c_{M,4}}\right)^2 (Mh)\left(c_{M,8}-c_{M,4}^2\right) R_8 \\
&= h\left(\frac{2M}{M+2}\right)^2 \frac{(M+2)(M+4)(M+6)-M(M+2)^2}{M^2} R_8 \\
&= O_{Q_h}(Mh) \\
&= o_{Q_h}(1) \text{ as } Mh \to 0.
\end{aligned}$$

Finally results follow in probability under P, by using Theorem 3.2.1.

Proof of Proposition 3.5.1 This follows from Theorem 3.11.1 and 3.11.2, given that conditionally on σ for any $q > 0$, $\frac{1}{c_{M,q}} R_q \to \overline{\sigma^q}$ in probability under Q_h and P (see Section 4.1 of Myklang and Zhang (2009)). For any $p, q > 1$, $\lim_{M\to\infty} \frac{c_{M,q}}{\left(c_{M,p}\right)^{q/p}} = 1$.

3.11.3 Appendix G

This appendix concerns the multivariate case where the parameter of interest is the integrated beta.

Notation

We introduce some notation.

$$T_{\beta,h,M} = S_{\beta,h,M}\left(\frac{\hat{V}_\beta}{V_{\beta,h,M}}\right)^{-1/2} = S_{\beta,h,M}\left(1+\sqrt{h}U_{h,M}\right)^{-1/2},$$

where

$$S_{\beta,h,M} = \frac{\sqrt{h^{-1}}\left(\hat{\beta}_{lk}-\beta_{lk}\right)}{\sqrt{V_{\beta,h,M}}} \text{ and } U_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\left(\hat{V}_{\beta,h,M}-V_{\beta,h,M}\right)}{V_{\beta,h,M}},$$

and $V_{\beta,h,M} \equiv Var\left(\sqrt{h^{-1}}\left(\hat{\beta}_{lk}-\beta_{lk}\right)\right) = \frac{M}{M-2}\sum_{j=1}^{1/Mh} Mh\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)$. We let

$$U_{1,\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\sum_{j=1}^{1/Mh}\left(V_{1,(j)}-E\left(V_{1,(j)}\right)\right)}{V_{\beta,h,M}} \text{ and } U_{2,\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\sum_{j=1}^{1/Mh}\left(V_{2,(j)}-E\left(V_{2,(j)}\right)\right)}{V_{\beta,h,M}},$$

where

$$V_{1,(j)} = \frac{M^2h}{M-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right), \text{ and}$$

$$V_{2,(j)} = \frac{M^2 h}{M-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^2.$$

We also let for any $q > M$, $R_{\beta,q} \equiv Mh \sum_{j=1}^{1/Mh}\left(\frac{M}{M-1}\right)^{\frac{q}{2}} \frac{1}{b_{M,q}c_{M-1,q}}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{q}{2}}$, where the definition of $c_{M,q}$ is given in equation (3.6), and for any $q > M$, we have $b_{M,q} \equiv E\left(\left(\frac{M}{\chi_M^2}\right)^{\frac{q}{2}}\right) = \left(\frac{M}{2}\right)^{\frac{q}{2}} \frac{\Gamma\left(\frac{M}{2}-\frac{q}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}$, where χ_M^2 is the standard χ^2 distribution with M degrees of freedom. Note that $b_{M,2} = \frac{M}{M-2}$, $b_{M,4} = \frac{M^2}{(M-2)(M-4)}$, and $b_{M,6} = \frac{M^3}{(M-2)(M-4)(M-6)}$. It follows by using the definition of $b_{M,q}$ and this property of the Gamma function, for all $x > 0$, $\Gamma(x+1) = x\Gamma(x)$. Finally we denote by $y_{k(j)} = \left(y_{k,1+(j-1)M}, \cdots, y_{k,Mj}\right)'$, the M returns of asset k observed within the block j.

Similarly for the bootstrap, we let $T_{\beta,h,M}^* = S_{\beta,h,M}^*\left(1+\sqrt{h}U_{\beta,h,M}^*\right)^{-1/2}$, where $S_{\beta,h,M}^* = \frac{\sqrt{h^{-1}}(\hat{\beta}_{lk}^* - \hat{\beta}_{lk})}{\sqrt{V_{\beta}^*}}$, $U_{\beta,h,M}^* \equiv \frac{\sqrt{h^{-1}}\left(\hat{V}_{\beta,h,M}^* - V_{\beta,h,M}^*\right)}{V_{\beta,h,M}^*}$ and $V_{\beta,h,M}^* = Var^*\left(\sqrt{h^{-1}}(\hat{\beta}_{lk}^* - \hat{\beta}_{lk})\right)$. We also let

$$U_{1,\beta,h,M}^* \equiv \frac{\sqrt{h^{-1}}\sum_{j=1}^{1/Mh}\left(V_{1,(j)}^* - E^*\left(V_{1,(j)}^*\right)\right)}{V_{\beta,h,M}^*} \text{ and } U_{2,\beta,h,M}^* \equiv \frac{\sqrt{h^{-1}}\sum_{j=1}^{1/Mh}\left(V_{2,(j)}^* - E\left(V_{2,(j)}^*\right)\right)}{V_{\beta,h,M}^*}$$

where

$V_{1,(j)}^* \equiv \frac{M^2 h}{M-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^{*2}\right)^{-1}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^{*2}\right)$ and

$V_{2,(j)}^* \equiv \frac{M^2 h}{M-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^{*2}\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^{*}u_{i+(j-1)M}^{*}\right)^2$.

Finally we let $y_{k(j)}^* = \left(y_{k,1+(j-1)M}^*, \cdots, y_{k,Mj}^*\right)'$.

Auxiliary Lemmas

Lemma 3.11.3. *Suppose (3.1) and (3.2) hold. Then, we have that*

$$\hat{V}_{\beta,h,M} = \sum_{j=1}^{1/Mh} V_{1,(j)} - \sum_{j=1}^{1/Mh} V_{2,(j)}.$$

Lemma 3.11.4. *Suppose (3.1) and (3.2) hold with W independent of Σ. Then, conditionally on Σ, and under Q_h, we have for any integer M such that $M \approx ch^{-\alpha}$ with $\alpha \in [0, 1/2)$,*

a1) $E\left(V_{1,(j)}\right) = \frac{M^3 h}{(M-1)(M-2)}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)$, *for $M > 2$;*

a2) $E\left(V_{1,(j)}^2\right) = \frac{M^5(M+2)}{(M-1)^2(M-2)(M-4)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$;*

a3) $E\left(V_{2,(j)}\right) = \frac{M^2 h}{(M-1)(M-2)}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)$, *for* $M > 2$*;*

a4) $E\left(V_{2,(j)}^2\right) = \frac{3M^4}{(M-1)^2(M-2)(M-4)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for* $M > 4$*;*

a5) $E\left(V_{1,(j)}V_{2,(j)}\right) = \frac{M^4(M+2)}{(M-1)^2(M-2)(M-4)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for* $M > 4$*;*

a6) $Var\left(V_{1,(j)}\right) = \frac{4M^5}{(M-1)(M-2)^2(M-4)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for* $M > 4$*;*

a7) $Var\left(V_{2,(j)}\right) = \frac{2M^4}{(M-1)(M-2)^2(M-4)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for* $M > 4$*;*

a8) $Cov\left(V_{1,(j)}, V_{2,(j)}\right) = \frac{4M^4}{(M-1)(M-2)^2(M-4)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for* $M > 4$*;*

a9) $Var\left(V_{1,(j)} - V_{2,(j)}\right) = \frac{2M^5(2M-3)}{(M-1)(M-2)^2(M-4)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for* $M > 4$.

Lemma 3.11.5. *Suppose (3.1) and (3.2) hold with* W *independent of* Σ*. Then, conditionally on* Σ*, and under* Q_h*, let* $M > 4$ *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$*,*

a1) $E\left(\hat{V}_{\beta,h,M}\right) = V_{\beta,h,M}$*;*

a2) $Var\left(\hat{V}_{\beta,h,M}\right) = \frac{2M^4(2M-3)}{(M-1)(M-2)^2(M-4)}h\left(Mh\sum_{j=1}^{1/Mh}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2\right)$*;*

a3) $\hat{V}_{\beta,h,M} - V_{\beta,h,M} \to 0$ *in probability;*

a4) $V_{\beta,h,M} \to V_\beta$.

Lemma 3.11.6. *Suppose (3.1) and (3.2) hold with* W *independent of* Σ*. Then, conditionally on* Σ*, and under* Q_h*, we have for any integer* M *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$*,*

a1) $E\left(S_{\beta,h,M}\right) = 0$*;*

a2) $E\left(S_{\beta,h,M}^2\right) = 1$*;*

a3) $E\left(S_{\beta,h,M}^3\right) = 0$*;*

a4) $E\left(S_{\beta,h,M}U_{1,\beta,h,M}\right) = 0$*;*

a5) $E\left(S_{\beta,h,M}U_{2,\beta,h,M}\right) = 0$*;*

a6) $E\left(S^{3}_{\beta,h,M}U_{1,\beta,h,M}\right)=0$;

a7) $E\left(S^{3}_{\beta,h,M}U_{2,\beta,h,M}\right)=0$.

Lemma 3.11.7. *Suppose (3.1) and (3.2) hold with W independent of Σ. Then, conditionally on Σ, and under Q_h, we have for any integer M such that $M\approx ch^{-\alpha}$ with $\alpha\in[0,1/2)$,*

a1) $(\hat{\Gamma}_{k(j)})^{-1}\sum_{i=1}^{M}\hat{u}^{2}_{i+(j-1)M}=\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^{2}$;

a2) $E\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^{2}\right)^{\frac{q}{2}}=\left(\frac{M-1}{M}\right)^{\frac{q}{2}}b_{M,q}c_{M-1,q}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^{2}\right)^{\frac{q}{2}}$, *for $M>q$;*

a3) $R_{\beta,q}\equiv Mh\sum_{j=1}^{1/Mh}\left(\frac{M^{\frac{q}{2}}}{(M-1)^{\frac{q}{2}}b_{M,q}c_{M-1,q}}\right)\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^{2}\right)^{\frac{q}{2}}-Mh\sum_{j=1}^{1/Mh}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^{2}\right)^{\frac{q}{2}}\to 0$ *in probability under Q_h and P, for any $M>q\,(1+\delta)$, for some $\delta>0$;*

a4) $\hat{V}_{\beta,h,M}-V_{\beta,h,M}\to 0$ *in probability under Q_h and P, for any $M>2\,(1+\delta)$, for some $\delta>0$.*

Proof of Lemma 3.11.3. Using the definition of $\hat{V}_{\beta,h,M}$ in the text (see Equation (3.16)), and the definition of $\hat{u}_{i+(j-1)M}=y_{l,i+(j-1)M}-\hat{\beta}_{lk(j)}y_{k,i+(j-1)M}$, we can write

$$\begin{aligned}\hat{V}_{\beta,h,M}&=M^{2}h\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y^{2}_{k,i+(j-1)M}\right)^{-1}\left(\frac{1}{M-1}\sum_{i=1}^{M}\left(y_{l,i+(j-1)M}-\hat{\beta}_{lk(j)}y_{k,i+(j-1)M}\right)^{2}\right)\\&=\frac{M^{2}h}{M-1}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y^{2}_{k,i+(j-1)M}\right)^{-1}\left(\sum_{i=1}^{M}\left(u_{i+(j-1)M}-\left(\hat{\beta}_{lk(j)}-\beta_{lk(j)}\right)y_{k,i+(j-1)M}\right)^{2}\right),\end{aligned}$$

where we used the definition of $y_{l,i+(j-1)M}$ see Equation (3.16). Adding and subtracting appropriately, it follows that

$$\begin{aligned}\hat{V}_{\beta,h,M}&=\frac{M^{2}h}{M-1}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y^{2}_{k,i+(j-1)M}\right)^{-1}\left(\left(\sum_{i=1}^{M}u^{2}_{i+(j-1)M}\right)+\left(\hat{\beta}_{lk(j)}-\beta_{lk(j)}\right)^{2}\right)\\&-2\frac{M^{2}h}{M-1}\sum_{j=1}^{1/Mh}\left(\hat{\beta}_{lk(j)}-\beta_{lk(j)}\right)\left(\sum_{i=1}^{M}y^{2}_{k,i+(j-1)M}\right)^{-1}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)\\&=\frac{M^{2}h}{M-1}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y^{2}_{k,i+(j-1)M}\right)^{-1}\left(\sum_{i=1}^{M}u^{2}_{i+(j-1)M}\right)\\&-\frac{M^{2}h}{M-1}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y^{2}_{k,i+(j-1)M}\right)^{-2}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^{2}\\&=\sum_{j=1}^{1/Mh}V_{1,(j)}-\sum_{j=1}^{1/Mh}V_{2,(j)},\end{aligned}$$

where we used $\left(\hat{\beta}_{lk(j)} - \beta_{lk(j)}\right) = \left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1} \left(\sum_{i=1}^{M} y_{k,i+(j-1)M} u_{i+(j-1)M}\right)$.

Proof of Lemma 3.11.4 part a1). Giving the definition of $V_{1,(j)}$, the law of iterated expectations and the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, we can write

$$\begin{aligned}
E\left(V_{1,(j)}\right) &= E\left(E\left(V_{1,(j)}|y_{k(j)}\right)\right) \\
&= \frac{M^2 h}{M-1} E\left(E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right)\right)|y_{k(j)}\right) \\
&= \frac{M^2 h}{M-1} E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} E\left(u_{i+(j-1)M}^2|y_{k(j)}\right)\right)\right) \\
&= \frac{M^3 h}{M-1} V_{(j)} E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\right),
\end{aligned}$$

then given equation (3.14) in the text and by replacing $V_{(j)}$ by $\frac{1}{M}\left(\Gamma_{l(j)} - \frac{\Gamma_{lk(j)}^2}{\Gamma_{k(j)}}\right)$, we have that

$$E\left(V_{1,(j)}\right) = \frac{M^3 h}{(M-1)(M-2)}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right).$$

Proof of Lemma 3.11.4 part a2). Given the definition of $V_{1,(j)}$ and the law of iterated expectations, we can write

$$\begin{aligned}
E\left(V_{1,(j)}^2\right) &= E\left(E\left(V_{1,(j)}^2|y_{k(j)}\right)\right) \\
&= \frac{M^4 h^2}{(M-1)^2} E\left(E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right)^2\right)|y_{k(j)}\right) \\
&= \frac{M^4 h}{(M-1)^2} V_{(j)}^2 E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2} E\left(\sum_{i=1}^{M}\left(\frac{u_{i+(j-1)M}}{\sqrt{V_{(j)}}}\right)^2 |y_{k(j)}\right)^2\right).
\end{aligned}$$

Note that since $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, $E\left(\sum_{i=1}^{M}\left(\frac{u_{i+(j-1)M}}{\sqrt{V_{(j)}}}\right)^2 |y_{k(j)}\right)^2 = E\left(\chi_{j,M}^2\right)^2 = M(M+2)$ where $\chi_{j,M}^2$ follow the standard χ^2 distribution with M degrees of freedom. Then we have

$$E\left(V_{1,(j)}^2\right) = \frac{M^5 (M+2) h}{(M-1)^2} V_{(j)}^2 E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\right),$$

then given the fact that $\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 \overset{d}{=} \frac{\Gamma_{k(j)}}{M}\chi_{j,M}^2$, where '$\overset{d}{=}$' denotes equivalence in distribution, by using the second moment of an inverse of χ^2 distribution, we have

$E\left(\frac{1}{\chi^2_{j,M}}\right)^2 = \frac{1}{(M-2)(M-4)}$, and by replacing $V_{(j)}$ by $\frac{1}{M}\left(\Gamma_{l(j)} - \frac{\Gamma^2_{lk(j)}}{\Gamma_{k(j)}}\right)$ it follows that

$$E\left(V^2_{1,(j)}\right) = \frac{M^5\,(M+2)}{(M-1)^2\,(M-2)\,(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2.$$

Proof of Lemma 3.11.4 part a3). Given the definition of $V_{1,(j)}$, the law of iterated expectations and the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$. we can write

$$\begin{aligned}
E\left(V_{2,(j)}\right) &= E\left(E\left(V_{2,(j)}|y_{k(j)}\right)\right) \\
&= \frac{M^2 h}{M-1} E\left(E\left(\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^2\right)|y_{k(j)}\right) \\
&= \frac{M^2 h}{M-1} E\left(\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-2}\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M} E\left(u^2_{i+(j-1)M}|y_{k(j)}\right)\right)\right) \\
&= \frac{M^2 h}{M-1} V_{(j)} E\left(\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-1}\right),
\end{aligned}$$

then using equation (3.14) in the text and replacing $V_{(j)}$ by $\frac{1}{M}\left(\Gamma_{l(j)} - \frac{\Gamma^2_{lk(j)}}{\Gamma_{k(j)}}\right)$ yields

$$E\left(V_{2,(j)}\right) = \frac{M^2 h}{(M-1)\,(M-2)} \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right).$$

Proof of Lemma 3.11.4 part a4). Given the definition of $V_{2,(j)}$ and the law of iterated expectations, we can write

$$\begin{aligned}
E\left(V^2_{2,(j)}\right) &= E\left(E\left(V^2_{2,(j)}|y_{k(j)}\right)\right) \\
&= \frac{M^4 h^2}{(M-1)^2} E\left(\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-4} E\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^4 |y_{k(j)}\right) \\
&\equiv \frac{M^4 h^2}{(M-1)^2} E\left(\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-4} A\right).
\end{aligned}$$

Then using the conditional independence and mean zero property of $y_{k,i+(j-1)M}u_{i+(j-1)M}$ we have that

$$
\begin{aligned}
A &\equiv E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^4 |y_{k(j)}\right)\\
&= \sum_{i=1}^{M} E\left(y^4_{k,i+(j-1)M}u^4_{i+(j-1)M}|y_{k(j)}\right)\\
&+ 3\sum_{i\neq s} E\left(y^2_{k,i+(j-1)M}u^2_{i+(j-1)M}|y_{k(j)}\right) E\left(y^2_{k,s+(j-1)M}u^2_{s+(j-1)M}|y_{k(j)}\right)\\
&= 3V^2_{(j)}\left(\sum_{i=1}^{M} y^4_{k,i+(j-1)M} + \sum_{i\neq s} y^2_{k,i+(j-1)M}y^2_{k,s+(j-1)M}\right)\\
&= 3V^2_{(j)}\left(\sum_{i=1}^{M} y^4_{k,i+(j-1)M}\right)^2,
\end{aligned}
$$

thus we can write

$$
E\left(V^2_{2,(j)}\right) = \frac{M^4h^2}{(M-1)^2}3V^2_{(j)}E\left(\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-2}\right),
$$

result follows similarly where we use the same arguments as in the proof of Lemma 3.11.4 part a2).

Proof of Lemma 3.11.4 part a5). The proof follows similarly as parts a2) and a4) of Lemma 3.11.4 and therefore we omit the details.

Proof of Lemma 3.11.5 part a1). Given the definition of $\hat{V}_{\beta,h,M}$, $V_{1,(j)}$, $V_{2,(j)}$ and by using Lemma 3.11.3 and part 1 of Lemma 3.11.4, we can write

$$
\begin{aligned}
E\left(\hat{V}_{\beta,h,M}\right) &= E\left(\sum_{j=1}^{1/Mh} V_{1,(j)}\right) - E\left(\sum_{j=1}^{1/Mh} V_{2,(j)}\right)\\
&= \sum_{j=1}^{1/Mh} E\left(V_{1,(j)}\right) - \sum_{j=1}^{1/Mh} E\left(V_{1,(j)}\right)\\
&= \frac{M^3h}{(M-1)(M-2)}\sum_{j=1}^{1/Mh}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)\\
&- \frac{M^2h}{(M-1)(M-2)}\sum_{j=1}^{1/Mh}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)\\
&= \frac{M}{M-1}V_{\beta,h,M} - \frac{1}{M-1}V_{\beta,h,M}\\
&= V_{\beta,h,M}.
\end{aligned}
$$

Proof of Lemma 3.11.5 part a2). Given the definitions of $\hat{V}_{\beta,h,M}$, $V_{1,(j)}$, $V_{2,(j)}$ and Lemma 3.11.3, we can write

$$Var\left(\hat{V}_{\beta,h,M}\right) = Var\left(\sum_{j=1}^{1/Mh} V_{1,(j)}\right) + Var\left(\sum_{j=1}^{1/Mh} V_{2,(j)}\right) - 2Cov\left(\sum_{j=1}^{1/Mh} V_{1,(j)}, \sum_{j=1}^{1/Mh} V_{2,(j)}\right),$$

given the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, we have $V_{1,(j)}$ and $V_{2,(j)}$ are conditionally independent given $y_{k(j)}$, $V_{1,(j)}$ and $V_{2,(t)}$ are conditionally independent for all $t \neq j$ given $y_{k(j)}$. It follows that

$$\begin{aligned} Var\left(\hat{V}_{\beta,h,M}\right) &= \sum_{j=1}^{1/Mh}\left(E\left(V_{1,(j)}^2\right) - E\left(V_{1,(j)}\right)^2\right) + \left(E\left(V_{2,(j)}^2\right) - E\left(V_{2,(j)}\right)^2\right) \\ &- 2\sum_{j=1}^{1/Mh}\left(E\left(V_{2,(j)}V_{2,(j)}\right) - E\left(V_{1,(j)}\right)E\left(V_{2,(j)}\right)\right), \end{aligned}$$

finally results follow given Lemma 3.11.4.

Proof of Lemma 3.11.5 part a3). Results follow directly given Lemma 3.11.4 parts a1) and a2) since $E\left(\hat{V}_{\beta,h,M} - V_{\beta,h,M}\right) = 0$ and $Var\left(\hat{V}_{\beta,h,M} - V_{\beta,h,M}\right) \to 0$ as $h \to 0$ provide that $Mh \to 0$

Proof of Lemma 3.11.5 part a4). This result follows from the boundedness of $\Sigma_k(u)$, $\Sigma_l(u)$ and the Reimann integrable of $\Sigma_{kl}(u)$ for any $k, l = 1 \cdots d$.

Proof of Lemma 3.11.6 part a1). Given the definition of $S_{\beta,h,M}$ we can write

$$\begin{aligned} E\left(S_{\beta,h,M}\right) &= \frac{\sqrt{h^{-1}}M}{\sqrt{V_{\beta,h,M}}}\sum_{i=1}^{M} E\left(\left(\hat{\beta}_{lk(j)} - \beta_{lk}(j)\right)\right) \\ &= 0, \end{aligned}$$

where the last equality use the unbiased property of OLS estimator $\hat{\beta}_{lk(j)}$.

Proof of Lemma 3.11.6 part a2). Given the definitions of $S_{\beta,h,M}$ and $V_{\beta,h,M}$ we can rewrite

$$\begin{aligned} Var\left(S_{\beta,h,M}\right) &= \frac{1}{V_{\beta,h,M}} Var\left(\sqrt{h}\left(\hat{\beta}_{lk} - \beta_{lk}\right)\right) \\ &= 1. \end{aligned}$$

Proof of Lemma 3.11.6 part a3). Given the definition of $S_{\beta,h,M}$ and the fact that we can write

$$\sqrt{h}\left(\hat{\beta}_{lk} - \beta_{lk}\right) = M\sqrt{h}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right),$$

we have that

$$E\left(S_{\beta,h,M}^{3}\right)=\frac{M^{3}h^{3/2}}{V_{\beta,h,M}^{3/2}}E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-1}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)\right)^{3},$$

then given the fact that $u_{i+(j-1)M}|y_{k(j)}\sim i.i.d.N\left(0,V_{(j)}\right)$, we have that

$$\begin{aligned}
E\left(S_{\beta,h,M}^{3}\right) &= \frac{M^{3}h^{3/2}}{V_{\beta,h,M}^{3/2}}E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-3}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^{3}\right)\\
&= \frac{M^{3}h^{3/2}}{V_{\beta,h,M}^{3/2}}E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-3}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{3}E\left(u_{i+(j-1)M}^{3}|y_{k(j)}\right)\right)\right)\\
&= 0.
\end{aligned}$$

Proof of Lemma 3.11.6 part a4). We start the proof by introducing this notation, which is relevant only for part a4) of Lemma 3.11.6. We let $B=E\left(S_{\beta,h,M}U_{1,\beta,h,M}\right)$, then giving the definitions of $S_{\beta,h,M}$, $U_{1,\beta,h,M}$ and by using the fact that $u_{i+(j-1)M}|y_{k(j)}\sim i.i.d.N\left(0,V_{(j)}\right)$, we can write

$$\begin{aligned}
B &= \frac{M^{3}h}{(M-1)V_{\beta,h,M}^{3/2}}\sum_{j=1}^{1/Mh}E\left(\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-2}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)\left(\sum_{i=1}^{M}u_{i+(j-1)M}^{2}\right)\right)\\
&- \frac{M}{V_{\beta,h,M}^{3/2}}\sum_{j=1}^{1/Mh}E\left(V_{1,(j)}\right)E\left(\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-1}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)\right),
\end{aligned}$$

results follow then, by using the law of iterated expectations and again the fact that $u_{i+(j-1)M}|y_{k(j)}\sim i.i.d.N\left(0,V_{(j)}\right)$.

Proof of Lemma 3.11.3 part a5). Giving the definition of $S_{\beta,h,M}$, $U_{2,\beta,h,M}$ and by using the fact that $u_{i+(j-1)M}|y_{k(j)}\sim i.i.d.N\left(0,V_{(j)}\right)$, we can write

$$\begin{aligned}
E\left(S_{\beta,h,M}U_{2,\beta,h,M}\right) &= \frac{M^{3}h^{3/2}}{(M-1)V_{\beta,h,M}^{3/2}}E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-3}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^{3}\right)\\
&- \frac{M}{V_{\beta,h,M}^{3/2}}\sum_{j=1}^{1/Mh}E\left(V_{2,(j)}\right)E\left(\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-1}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)\right)
\end{aligned}$$

then results follow by using the law of iterated expectations and again the fact that $u_{i+(j-1)M}|y_{k(j)}\sim i.i.d.N\left(0,V_{(j)}\right)$.

The proof of the remaining results (Lemma 3.11.6 part a6) and part a7)) follow similarly and therefore we omit the details.

Proof of Lemma 3.11.7 part a1). Given the definition of $\hat{u}_{i+(j-1)M}$, we can write

$$
\begin{aligned}
(\hat{\Gamma}_{k(j)})^{-1}\sum_{i=1}^{M}\hat{u}_{i+(j-1)M}^{2} &= \frac{1}{\hat{\Gamma}_{k(j)}}\sum_{i=1}^{M}\left(y_{l,i+(j-1)M}-\hat{\beta}_{lk(j)}y_{k,i+(j-1)M}\right)^{2}\\
&= \frac{1}{\hat{\Gamma}_{k(j)}}\sum_{i=1}^{M}\left(y_{l,i+(j-1)M}^{2}-2\hat{\beta}_{lk(j)}y_{l,i+(j-1)M}y_{k,i+(j-1)M}+\hat{\beta}_{lk(j)}^{2}y_{k,i+(j-1)M}^{2}\right)\\
&= \frac{1}{\hat{\Gamma}_{k(j)}}\left(\sum_{i=1}^{M}y_{l,i+(j-1)M}^{2}-2\hat{\beta}_{lk(j)}\sum_{i=1}^{M}y_{l,i+(j-1)M}y_{k,i+(j-1)M}\right)\\
&+ \frac{1}{\hat{\Gamma}_{k(j)}}\hat{\beta}_{lk(j)}^{2}\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2},
\end{aligned}
$$

thus results follow by replacing $\hat{\beta}_{lk(j)}=\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}$.

Proof of Lemma 3.11.7 part a2). Given the definitions of $\hat{\Gamma}_{l(j)}$, $\hat{\Gamma}_{k(j)}$ and $\hat{\Gamma}_{lk(j)}$ and by using part a1) of Lemma 3.11.7, we can write

$$
\begin{aligned}
E\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^{2}\right)^{\frac{q}{2}} &= E\left(\frac{\sum_{i=1}^{M}\hat{u}_{i+(j-1)M}^{2}}{\hat{\Gamma}_{k(j)}}\right)^{\frac{q}{2}}\\
&= E\left(\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-\frac{q}{2}}E\left(\sum_{i=1}^{M}\hat{u}_{i+(j-1)M}^{2}\right)^{\frac{q}{2}}|y_{k(j)}\right)\\
&= E\left(\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-\frac{q}{2}}V_{(j)}^{\frac{q}{2}}E\left(\left(\sum_{i=1}^{M}\frac{\hat{u}_{i+(j-1)M}^{2}}{V_{(j)}}\right)^{\frac{q}{2}}|y_{k(j)}\right)\right),
\end{aligned}
$$

where we use the law of iterated expectations and the fact that $u_{i+(j-1)M}|y_{k(j)}\sim i.i.d.N\left(0,V_{(j)}\right)$. Then given the definition of $c_{M,q}$ we can write

$$
\begin{aligned}
E\left(\left(\sum_{i=1}^{M}\frac{\hat{u}_{i+(j-1)M}^{2}}{V_{(j)}}\right)^{\frac{q}{2}}|y_{k(j)}\right) &= E\left(\left(\chi_{j,M}^{2}\right)^{\frac{q}{2}}\right)\\
&= (M-1)^{\frac{q}{2}}c_{M-1,q},
\end{aligned}
$$

it follows then that,

$$
\begin{aligned}
E\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^{2}\right)^{\frac{q}{2}} &= E\left(\frac{\sum_{i=1}^{M}\hat{u}_{i+(j-1)M}^{2}}{\hat{\Gamma}_{k(j)}}\right)^{\frac{q}{2}} \\
&= (M-1)^{\frac{q}{2}}c_{M-1,q}V_{(j)}^{\frac{q}{2}}E\left(\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2}\right)^{-\frac{q}{2}}\right) \\
&= (M-1)^{\frac{q}{2}}c_{M-1,q}V_{(j)}^{\frac{q}{2}}\Gamma_{k(j)}^{-\frac{q}{2}}E\left(\left(\frac{M}{\chi_{j,M}^{2}}\right)^{\frac{q}{2}}\right) \\
&= \left(\frac{M-1}{M}\right)^{\frac{q}{2}}b_{M,q}c_{M-1,q}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^{2}\right)^{\frac{q}{2}};
\end{aligned}
$$

where $b_{M,q}=E\left(\left(\frac{M}{\chi_{j,M}^{2}}\right)^{\frac{q}{2}}\right)$, for $M>q$.

Proof of Lemma 3.11.7 part a3). We verify the moments conditions of the Weak Law of Large Numbers for independent and nonidentically distributed on z_j, $j=1,\ldots,\frac{1}{Mh}$ with $z_j\equiv\frac{M^{\frac{q}{2}}}{(M-1)^{\frac{q}{2}}b_{M,q}c_{M-1,q}}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^{2}\right)^{\frac{q}{2}}$. By using part a2) of Lemma 3.11.7, for any $\delta>0$, and conditionally on σ, we can write

$$
E\left|z_j\right|^{1+\delta}=\left(\frac{M-1}{M}\right)^{\frac{\delta q}{2}}\frac{b_{M,(1+\delta)q}}{b_{M,q}}\frac{c_{M-1,(1+\delta)q}}{c_{M-1,q}}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^{2}\right)^{\frac{(1+\delta)q}{2}}<\infty
$$

since Σ is an adapted càdlàg spot covolatility matrix and locally bounded and invertible (in particular, $\Gamma_{k(j)}>0$), moreover in the case where $M\to\infty$, as $h\to 0$ (i.e. $M\approx ch^{-\alpha}$ with $\alpha\in(0,1/2)$) we have $\left(\frac{M-1}{M}\right)^{\frac{\delta q}{2}}\frac{b_{M,(1+\delta)q}}{b_{M,q}}\frac{c_{M-1,(1+\delta)q}}{c_{M-1,q}}\to 1$.

Proof of Lemma 3.11.7 part a4). Result follows directly given the definition of $\hat{V}_{\beta,h,M}, V_{\beta,h,M}$ and part a3) of Lemma 3.11.7, where we let $q=2$.

Remark 5 The bootstrap analogue of Lemma 3.11.3 and 3.11.4 replace $V_{1(j)}$ with $V_{1(j)}^{*}$, $V_{2(j)}$, the bootstrap analogue of Lemma 3.11.5 replaces $\hat{V}_{\beta,h,M}$ with $\hat{V}_{\beta,h,M}^{*}$, $V_{\beta,h,M}$ with $V_{\beta,h,M}^{*}$, $\Gamma_{l(j)}$ with $\hat{\Gamma}_{l(j)}$, $\Gamma_{k(j)}$ with $\hat{\Gamma}_{k(j)}$, and $\Gamma_{lk(j)}$ with $\hat{\Gamma}_{lk(j)}$; whereas the bootstrap analogue of Lemma 3.11.6 replaces $S_{\beta,h,M}$ with $S_{\beta,h,M}^{*}$, $U_{1,\beta,h,M}$ with $U_{1,\beta,h,M}^{*}$ and $U_{2,\beta,h,M}$ with $U_{2,\beta,h,M}^{*}$.

Lemma 3.11.8. *Suppose (3.1) and (3.2) hold with W independent of Σ. Then, conditionally on Σ, we have for any integer M such that $M\approx ch^{-\alpha}$ with $\alpha\in[0,1/2)$, and for some small $\delta>0$,*

a1) $E^{*}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^{2*}\right)^{-2(2+\delta)}=b_{M,4(2+\delta)}\hat{\Gamma}_{k(j)}^{-2(2+\delta)}$, *for $M>4\,(2+\delta)$;*

a2) $E^*\left(\left|\sum_{i=1}^{M} y^*_{k,i+(j-1)M} u^*_{i+(j-1)M}\right|^{2(2+\delta)}\right) \leq \mu^2_{2(2+\delta)} M^{2+\delta} \hat{\Gamma}^{2+\delta}_{k(j)} \hat{V}^{2+\delta}_{(j)}$;

Proof of Lemma 3.11.8 part a1). Given the definition of $y^*_{k,i+(j-1)M}$, we can write, $\sum_{i=1}^{M} y^{2*}_{k,i+(j-1)M} \overset{d}{=} \frac{\hat{\Gamma}_{k(j)}}{M} \sum_{i=1}^{M} v^2_{i+(j-1)M} \overset{d}{=} \frac{\hat{\Gamma}_{k(j)}}{M} \chi^2_{j,M}$, where $v_{i+(j-1)M} \sim i.i.d.N(0,1)$, and $\chi^2_{j,M}$ follow the standard χ^2 distribution with M degrees of freedom. Then for any integer $M > 4(2+\delta)$, we have that,

$$E\left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M}\right)^{-2(2+\delta)} = E\left(\frac{M}{\chi^2_{j,M}}\right)^{2(2+\delta)} \hat{\Gamma}^{-2(2+\delta)}_{k(j)} = b_{M,4(2+\delta)} \hat{\Gamma}^{-2(2+\delta)}_{k(j)}.$$

Proof of Lemma 3.11.8 part a2). Indeed by using the C_r inequality, and the law of iterated expectations and the fact that $u^*_{i+(j-1)M} | y^*_{k(j)} \sim i.i.d.N\left(0, \hat{V}_{(j)}\right)$. We can write for any $\delta > 0$,

$$\begin{aligned} E^*\left(\left|\sum_{i=1}^{M} y^*_{k,i+(j-1)M} u^*_{i+(j-1)M}\right|^{2(2+\delta)}\right) &\leq M^{3+2\delta} \sum_{i=1}^{M} E^*\left|y^*_{k,i+(j-1)M} u^*_{i+(j-1)M}\right|^{2(2+\delta)} \\ &= M^{3+2\delta} \sum_{i=1}^{M} E^*\left(y^{*2(2+\delta)}_{k,i+(j-1)M} E^*\left(u^{*2(2+\delta)}_{i+(j-1)M} | y^*_{k(j)}\right)\right) \\ &= \mu^2_{2(2+\delta)} M^{2+\delta} \hat{\Gamma}^{2+\delta}_{k(j)} \hat{V}^{2+\delta}_{(j)}, \end{aligned}$$

where the last equality follows since $y^{2*}_{k,i+(j-1)M} \overset{d}{=} \frac{\Gamma_{k(j)}}{M} v^2_{i+(j-1)M}$, where $v_{i+(j-1)M} \sim i.i.d.N(0,1)$ and $\mu_{2(2+\delta)} = E|v|^{2(2+\delta)}$.

Proof of proposition 3.5.2. As in Theorem 3.11.1, the first and third cumulants of $T_{\beta,h,M}$ are given by

$$\begin{aligned} \kappa_1(T_{\beta,h,M}) &= E(T_{\beta,h,M}), \\ \kappa_3(T_{\beta,h,M}) &= E(T^3_{\beta,h,M}) - 3E(T^2_{\beta,h,M})E(T_{\beta,h,M}) + 2[E(T_{\beta,h,M})]^3. \end{aligned}$$

Here, our goal is to identify the terms of order up to $O(\sqrt{h})$ of the asymptotic expansions of these two cumulants. We will first provide asymptotic expansions through order $O(\sqrt{h})$ for the first three moments of $T_{\beta,h,M}$. Note that for a given fixed value of k, a first-order Taylor expansion of $f(x) = (1+x)^{-k/2}$ around 0 yields $f(x) = 1 - \frac{k}{2}x + O(x^2)$. We have for any fixed integer k, We have that for any fixed integer k,

$$\begin{aligned} T^k_{\beta,h,M} &= S^k_{\beta,h,M}\left(1+\sqrt{h}U_{\beta,h,M}\right)^{-k/2}, \\ &= S^k_{\beta,h,M} - \frac{k}{2}\sqrt{h}S^k_{\beta,h,M}U_{\beta,h,M} + O(h) \\ &= S^k_{\beta,h,M} - \frac{k}{2}\sqrt{h}S^k_{\beta,h,M}U_{1,\beta,h,M} + \frac{k}{2}\sqrt{h}S^k_{\beta,h,M}U_{2,\beta,h,M} + O(h). \end{aligned}$$

For $k = 1, \cdots, 3$, the moments of $T^k_{\beta,h,M}$ up to order $O(h)$ are given by

$$\begin{aligned}
E\left(T_{\beta,h,M}\right) &= E\left(S_{\beta,h,M}\right) - \frac{\sqrt{h}}{2} E\left(S_{\beta,h,M} U_{1,\beta,h,M}\right) + \frac{\sqrt{h}}{2} E\left(S_{\beta,h,M} U_{2,\beta,h,M}\right) \\
E\left(T^2_{\beta,h,M}\right) &= E\left(S^2_{\beta,h,M}\right) - \sqrt{h} E\left(S^2_{\beta,h,M} U_{1,\beta,h,M}\right) + \sqrt{h} E\left(S^2_{\beta,h,M} U_{2,\beta,h,M}\right) \\
E\left(T^3_{\beta,h,M}\right) &= E\left(S^3_{\beta,h,M}\right) - \sqrt{h}\frac{3}{2} E\left(S^3_{\beta,h,M} U_{1,\beta,h,M}\right) + \sqrt{h}\frac{3}{2} E\left(S^3_{\beta,h,M} U_{2,\beta,h,M}\right).
\end{aligned}$$

Given Lemma 3.11.6, we have that

$$\begin{aligned}
E\left(T_{\beta,h,M}\right) &= 0 \\
E\left(T^2_{\beta,h,M}\right) &= 1 - \sqrt{h} E\left(S^2_{\beta,h,M} U_{1,\beta,h,M}\right) + \sqrt{h} E\left(S^2_{\beta,h,M} U_{2,\beta,h,M}\right) \\
E\left(T^3_{\beta,h,M}\right) &= 0.
\end{aligned}$$

It follows that $\kappa_1(T_{\beta,h,M}) = 0$ and $\kappa_3(T_{\beta,h,M}) = 0$.

Proof of Theorem 3.4.1 For part a), the proof follows the same steps as the proof of $V_{\beta,h,M}$ which we explain in the main text, in particular, given the definition of $\hat{\beta}^*_{lk}$, we have that

$$\begin{aligned}
V^*_{\beta,h,M} &= Var^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})\right) \\
&= M^2 h \sum_{j=1}^{1/Mh} Var^*\left(\hat{\beta}^*_{lk(j)} - \hat{\beta}_{lk(j)}\right) \\
&= M^2 h \sum_{j=1}^{1/Mh} E^*\left(\left(\sum_{i=1}^{M} y^{2*}_{k,i+(j-1)M}\right)^{-1}\right) \hat{V}_{(j)} \\
&= \frac{M^2 h}{M-2} \sum_{j=1}^{1/Mh} \left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right) \\
&= \frac{M-1}{M-2} \hat{V}_{\beta,h,M},
\end{aligned}$$

then results follows, given Lemma 3.11.5 or part a4) of Lemma 3.11.7.

For part b), we have $\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk}) = \sum_{j=1}^{1/Mh} z^*_{j,\beta}$, where

$$z^*_{j,\beta} = M\sqrt{h}\left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M}\right)^{-1}\left(\sum_{i=1}^{M} y^*_{k,i+(j-1)M} u^*_{i+(j-1)M}\right).$$

Note that $E^*\left(z^*_{j,\beta}\right) = 0$, and that

$$Var^*\left(\sum_{j=1}^{1/Mh} z^*_j\right) = V^*_{\beta,h,M} \overset{P}{\to} V_\beta,$$

by part a) moreover, since $z^*_1, \ldots, z^*_{1/Mh}$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta > 0$ and for some constant $C > 0$,

$$\sup_{x\in\Re}\left|P^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk}) \leq x\right) - \Phi\left(\frac{x}{V_\beta}\right)\right| \leq C\sum_{j=1}^{1/Mh} E^*\left|z^*_j\right|^{2+\delta},$$

Next, we show that $\sum\limits_{j=1}^{1/Mh} E^*\left|z^*_j\right|^{2+\delta} = o_p(1)$. We have that

$$\begin{aligned}
\sum_{j=1}^{1/Mh} E^*\left|z^*_{j,\beta}\right|^{2+\delta} &= \left(M\sqrt{h}\right)^{2+\delta}\sum_{j=1}^{1/Mh} E^*\left(\left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M}\right)^{-(2+\delta)}\left|\sum_{i=1}^{M} y^*_{k,i+(j-1)M}u^*_{i+(j-1)M}\right|^{2+\delta}\right) \\
&\equiv \left(M\sqrt{h}\right)^{2+\delta}\sum_{j=1}^{1/Mh} E^*\left(A^*_j B^*_j\right),
\end{aligned}$$

it follows then by using Cauchy-Schwarz inequality that

$$\begin{aligned}
E^*\left(A^*_j B^*_j\right) &\leq \sqrt{\left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M}\right)^{-2(2+\delta)}}\sqrt{E^*\left(\left|\sum_{i=1}^{M} y^*_{k,i+(j-1)M}u^*_{i+(j-1)M}\right|^{2(2+\delta)}\right)} \\
&\leq \mu_{2(2+\delta)} b^{\frac{1}{2}}_{M,4(2+\delta)} M^{1+\frac{\delta}{2}}\hat{\Gamma}^{-\frac{2+\delta}{2}}_{k(j)}\hat{V}^{\frac{2+\delta}{2}}_{(j)} \\
&= \mu_{2(2+\delta)} b^{\frac{1}{2}}_{M,4(2+\delta)}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{2+\delta}{2}},
\end{aligned}$$

where the second inequatily used part a1) and a2) of Lemma 3.11.8 and $\mu_{2(2+\delta)} = E\left|v\right|^{2(2+\delta)}$ with $v \sim N(0,1)$. Finally, given the definition of $R_{\beta,2+\delta}$ and the fact that

$M \approx ch^{-\alpha}$, we can write

$$\begin{aligned}
\sum_{j=1}^{1/Mh} E^* \left| z_{j,\beta}^* \right|^{2+\delta} &\leq \mu_{2(2+\delta)} b_{M,4(2+\delta)}^{\frac{1}{2}} M^{2+\delta} h^{1+\frac{\delta}{2}} \sum_{j=1}^{1/Mh} \left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}} \right)^2 \right)^{\frac{2+\delta}{2}} \\
&= \mu_{2(2+\delta)} b_{M,4(2+\delta)}^{\frac{1}{2}} \left(\frac{M-1}{M} \right)^{\frac{2+\delta}{2}} b_{M,2+\delta} c_{M-1,2+\delta} M^{1+\delta} h^{\frac{\delta}{2}} R_{\beta,2+\delta} \\
&= O_p \left(h^{\frac{\delta}{2} - \alpha(1+\delta)} b_{M,4(2+\delta)}^{\frac{1}{2}} \left(\frac{M-1}{M} \right)^{\frac{2+\delta}{2}} b_{M,2+\delta} c_{M-1,2+\delta} \right) \\
&= o_p(1).
\end{aligned}$$

Since for any $\delta > 0$, such that $\alpha \in [0, 1/2)$ we have $\frac{\delta}{2} - \alpha(1+\delta) > 0$, and $\mu_{2(2+\delta)} = E|v|^{2(2+\delta)} \leq \Delta < \infty$ where $v \sim N(0,1)$, moreover as $h \to 0$, $c_{M-1,2+\delta} \to 1$, $b_{M,4(2+\delta)} \to 1$, $b_{M,2+\delta} \to 1$ and by using Lemma 3.11.7 we have $R_{\beta,2+\delta} = O_P(1)$.

Proof of Theorem 3.4.2 Let

$$H_{\beta,h,M}^* = \frac{\sqrt{h^{-1}}(\hat{\beta}_{lk}^* - \hat{\beta}_{lk})}{\sqrt{V_{\beta,h,M}^*}},$$

and note that

$$T_{\beta,h,M}^* = H_{\beta,h,M}^* \sqrt{\frac{V_{\beta,h,M}^*}{\hat{V}_{\beta,h,M}^*}},$$

where $\hat{V}_{\beta,h,M}^*$ is defined in the main text. Theorem 3.4.1 proved that $H_{\beta,h,M}^* \xrightarrow{d^*} N(0,1)$ in probability. Thus, it suffices to show that $\hat{V}_{\beta,h,M}^* - V_{\beta,h,M}^* \xrightarrow{P^*} 0$ in probability under Q_h and P. In particular, we show that (1) $Bias^*\left(\hat{V}_{\beta,h,M}^*\right) = 0$, and (2) $Var^*\left(\hat{V}_{\beta,h,M}^*\right) \xrightarrow{P} 0$. Results follows directly by using the bootstrap analogue of parts a1), a2) and a3) of Lemma 3.11.5.